

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
18 April 2002 (18.04.2002)

PCT

(10) International Publication Number
WO 02/30268 A2

(51) International Patent Classification⁷: **A61B**

Brisbane, CA 94005 (US). **HEVEZI, Peter**; 1360 11th Avenue, San Francisco, CA 94122 (US).

(21) International Application Number: PCT/US01/32045

(22) International Filing Date: 12 October 2001 (12.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

09/687,576	13 October 2000 (13.10.2000)	US
09/733,742	8 December 2000 (08.12.2000)	US
09/733,288	8 December 2000 (08.12.2000)	US
60/263,957	24 January 2001 (24.01.2001)	US
60/276,888	16 March 2001 (16.03.2001)	US
60/276,791	16 March 2001 (16.03.2001)	US
60/281,922	6 April 2001 (06.04.2001)	US
60/286,214	24 April 2001 (24.04.2001)	US
09/847,046	30 April 2001 (30.04.2001)	US
60/288,589	4 May 2001 (04.05.2001)	US

(74) Agents: **BASTIAN, Kevin, L.** et al.; Townsend and Townsend and Crew LLP, Two Embarcadero Center, 8th Floor, San Francisco, CA 94111-3834 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(71) Applicant: **EOS BIOTECHNOLOGY, INC.** [US/US]; 225A Gateway Boulevard, South San Francisco, CA 94080-7019 (US).

(72) Inventors: **GISH, Kurt, C.**; 40 Perego Terrace #2, San Francisco, CA 94131 (US). **MACK, David, H.**; 2076 Monterey Avenue, Menlo Park, CA 94025 (US). **WILSON, Keith, E.**; 219 Jeter Street, Redwood City, CA 94062 (US). **AFAR, Daniel**; 435 Visitacion Avenue,

(54) Title: METHODS OF DIAGNOSIS OF PROSTATE CANCER, COMPOSITIONS AND METHODS OF SCREENING FOR MODULATORS OF PROSTATE CANCER

(57) Abstract: Described herein are genes whose expression are up-regulated or down-regulated in prostate cancer. Also described are such genes whose expression is further up-regulated or down-regulated in drug-resistant prostate cancer cells. Related methods and compositions that can be used for diagnosis and treatment of prostate cancer are disclosed. Also described herein are methods that can be used to identify modulators of prostate cancer.



WO 02/30268 A2

**METHODS OF DIAGNOSIS OF PROSTATE CANCER,
COMPOSITIONS AND METHODS OF SCREENING FOR
MODULATORS OF PROSTATE CANCER**

5

CROSS-REFERENCES TO RELATED APPLICATIONS

This application claims priority from the following applications: USSN 09/687,576 filed October 13, 2000, USSN 60/276,791 filed March 16, 2001; USSN 60/288,589, filed May 4, 2001; USSN 09/733,742, filed December 8, 2000; USSN 10 09/733,288, filed December 8, 2000; USSN 09/847,046, filed April 30, 2001; USSN 60/276,888, filed March 16, 2001; USSN 60/286,214, filed April 24, 2001; USSN 60/281,922, filed April 6, 2001; USSN 60/263,957, filed January 24, 2001, which are incorporated herein by reference in their entirety.

15

FIELD OF THE INVENTION

The invention relates to the identification of nucleic acid and protein expression profiles and nucleic acids, products, and antibodies thereto that are involved in prostate cancer; and to the use of such expression profiles and compositions in the diagnosis, prognosis and therapy of prostate cancer. The invention further relates to methods for 20 identifying and using agents and/or targets that inhibit prostate cancer.

BACKGROUND OF THE INVENTION

Prostate cancer is the most commonly diagnosed internal malignancy and second most common cause of cancer death in men in the U.S., resulting in approximately 25 40,000 deaths each year (Landis et al., *CA Cancer J. Clin.* 48:6-29 (1998); Greenlee et al., *CA Cancer J. Clin.* 50(1):7-13 (2000)), and incidence of prostate cancer has been increasing rapidly over the past 20 years in many parts of the world (Nakata et al., *Int. J. Urol.* 7(7):254-257 (2000); Majeed et al., *BJU Int.* 85(9):1058-1062 (2000)). It develops as the

result of a pathologic transformation of normal prostate cells. In tumorigenesis, the cancer cell undergoes initiation, proliferation and loss of contact inhibition, culminating in invasion of surrounding tissue and, ultimately, metastasis.

Deaths from prostate cancer are a result of metastasis of a prostate tumor.

5 Therefore, early detection of the development of prostate cancer is critical in reducing mortality from this disease. Measuring levels of prostate-specific antigen (PSA) has become a very common method for early detection and screening, and may have contributed to the slight decrease in the mortality rate from prostate cancer in recent years (Nowroozi et al., *Cancer Control* 5(6):522-531 (1998)). However, many cases are not diagnosed until the
10 disease has progressed to an advanced stage.

Treatments such as surgery (prostatectomy) , radiation therapy, and cryotherapy are potentially curative when the cancer remains localized to the prostate. Therefore, early detection of prostate cancer is important for a positive prognosis for treatment. Systemic treatment for metastatic prostate cancer is limited to hormone therapy
15 and chemotherapy. Chemical or surgical castration has been the primary treatment for symptomatic metastatic prostate cancer for over 50 years. This testicular androgen deprivation therapy usually results in stabilization or regression of the disease (in 80% of patients), but progression of metastatic prostate cancer eventually develops (Panvichian et al., *Cancer Control* 3(6):493-500 (1996)). Metastatic disease is currently considered incurable,
20 and the primary goals of treatment are to prolong survival and improve quality of life (Rago, *Cancer Control* 5(6):513-521 (1998)).

Thus, methods that can be used for diagnosis and prognosis of prostate cancer and effective treatment of prostate cancer, and including particularly metastatic prostate cancer, would be desirable. Accordingly, provided herein are methods that can be used in
25 diagnosis and prognosis of prostate cancer. Further provided are methods that can be used to screen candidate bioactive agents for the ability to modulate, e.g., treat, prostate cancer. Additionally, provided herein are molecular targets and compositions for therapeutic intervention in prostate cancer and other cancers.

SUMMARY OF THE INVENTION

The present invention therefore provides nucleotide sequences of genes that are up- and down-regulated in prostate cancer cells. Such genes are useful for diagnostic purposes, and also as targets for screening for therapeutic compounds that modulate prostate cancer, such as hormones or antibodies. Other aspects of the invention will become apparent to the skilled artisan by the following description of the invention.

In one aspect, the present invention provides a method of detecting a prostate cancer-associated transcript in a cell from a patient, the method comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-16.

In one embodiment, the present invention provides a method of determining the level of a prostate cancer associated transcript in a cell from a patient.

In one embodiment, the present invention provides a method of detecting a prostate cancer-associated transcript in a cell from a patient, the method comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-16.

In one embodiment, the polynucleotide selectively hybridizes to a sequence at least 95% identical to a sequence as shown in Tables 1-16. In another embodiment, the polynucleotide comprises a sequence as shown in Tables 1-16.

In one embodiment, the biological sample is a tissue sample. In another embodiment, the biological sample comprises isolated nucleic acids, e.g., mRNA.

In one embodiment, the polynucleotide is labeled, e.g., with a fluorescent label.

In one embodiment, the polynucleotide is immobilized on a solid surface.

In one embodiment, the patient is undergoing a therapeutic regimen to treat prostate cancer. In another embodiment, the patient is suspected of having metastatic prostate cancer.

In one embodiment, the patient is a human.

In one embodiment, the patient is suspected of having a taxol-resistant cancer.

In one embodiment, the prostate cancer associated transcript is mRNA.

In one embodiment, the method further comprises the step of amplifying nucleic acids before the step of contacting the biological sample with the polynucleotide.

In another aspect, the present invention provides a method of monitoring the efficacy of a therapeutic treatment of prostate cancer, the method comprising the steps of: (i) providing a biological sample from a patient undergoing the therapeutic treatment; and (ii) determining the level of a prostate cancer-associated transcript in the biological sample by contacting the biological sample with a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-16, thereby monitoring the efficacy of the therapy. In a further embodiment, the patient has metastatic prostate cancer. In a further embodiment, the patient has a drug resistant (e.g., taxol resistant) form of prostate cancer.

In one embodiment, the method further comprises the step of: (iii) comparing the level of the prostate cancer-associated transcript to a level of the prostate cancer-associated transcript in a biological sample from the patient prior to, or earlier in, the therapeutic treatment.

Additionally, provided herein is a method of evaluating the effect of a candidate prostate cancer drug comprising administering the drug to a patient and removing a cell sample from the patient. The expression profile of the cell is then determined. This method may further comprise comparing the expression profile to an expression profile of a healthy individual. In a preferred embodiment, said expression profile includes a gene of Tables 1-16.

In one aspect, the present invention provides an isolated nucleic acid molecule consisting of a polynucleotide sequence as shown in Tables 1-16.

In one embodiment, an expression vector or cell comprises the isolated nucleic acid.

In one aspect, the present invention provides an isolated polypeptide which is encoded by a nucleic acid molecule having polynucleotide sequence as shown in Tables 1-16.

In another aspect, the present invention provides an antibody that specifically binds to an isolated polypeptide which is encoded by a nucleic acid molecule having polynucleotide sequence as shown in Tables 1-16.

In one embodiment, the antibody is conjugated to an effector component, e.g., a fluorescent label, a radioisotope or a cytotoxic chemical.

In one embodiment, the antibody is an antibody fragment. In another embodiment, the antibody is humanized.

5 In one aspect, the present invention provides a method of detecting a prostate cancer cell in a biological sample from a patient, the method comprising contacting the biological sample with an antibody as described herein.

In another aspect, the present invention provides a method of detecting antibodies specific to prostate cancer in a patient, the method comprising contacting a
10 biological sample from the patient with a polypeptide encoded by a nucleic acid comprising a sequence from Tables 1-16.

In another aspect, the present invention provides a method for identifying a compound that modulates a prostate cancer-associated polypeptide, the method comprising the steps of: (i) contacting the compound with a prostate cancer-associated polypeptide, the
15 polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-16; and (ii) determining the functional effect of the compound upon the polypeptide.

In one embodiment, the functional effect is a physical effect, an enzymatic effect, or a chemical effect.

20 In one embodiment, the polypeptide is expressed in a eukaryotic host cell or cell membrane. In another embodiment, the polypeptide is recombinant.

In one embodiment, the functional effect is determined by measuring ligand binding to the polypeptide.

In another aspect, the present invention provides a method of inhibiting
25 proliferation of a prostate cancer-associated cell to treat prostate cancer in a patient, the method comprising the step of administering to the subject a therapeutically effective amount of a compound identified as described herein.

In one embodiment, the compound is an antibody.

In another aspect, the present invention provides a drug screening assay
30 comprising the steps of: (i) administering a test compound to a mammal having prostate cancer or to a cell sample isolated therefrom; (ii) comparing the level of gene expression of a

polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-16 in a treated cell or mammal with the level of gene expression of the polynucleotide in a control cell sample or mammal, wherein a test compound that modulates the level of expression of the polynucleotide is a candidate for the treatment of prostate cancer.

In one embodiment, the control is a mammal with prostate cancer or a cell sample therefrom that has not been treated with the test compound. In another embodiment, the control is a normal cell or mammal.

In one embodiment, the test compound is administered in varying amounts or concentrations. In another embodiment, the test compound is administered for varying time periods. In another embodiment, the comparison can occur after addition or removal of the drug candidate.

In one embodiment, the levels of a plurality of polynucleotides that selectively hybridize to a sequence at least 80% identical to a sequence as shown in Tables 1-16 are individually compared to their respective levels in a control cell sample or mammal. In a preferred embodiment the plurality of polynucleotides is from three to ten.

In another aspect, the present invention provides a method for treating a mammal having prostate cancer comprising administering a compound identified by the assay described herein.

In another aspect, the present invention provides a pharmaceutical composition for treating a mammal having prostate cancer, the composition comprising a compound identified by the assay described herein and a physiologically acceptable excipient.

In one aspect, the present invention provides a method of screening drug candidates by providing a cell expressing a gene that is up- and down-regulated as in a prostate cancer. In one embodiment, a gene is selected from Tables 1-16. The method further includes adding a drug candidate to the cell and determining the effect of the drug candidate on the expression of the expression profile gene.

In one embodiment, the method of screening drug candidates includes comparing the level of expression in the absence of the drug candidate to the level of expression in the presence of the drug candidate, wherein the concentration of the drug

candidate can vary when present, and wherein the comparison can occur after addition or removal of the drug candidate. In a preferred embodiment, the cell expresses at least two expression profile genes. The profile genes may show an increase or decrease.

Also provided is a method of evaluating the effect of a candidate prostate cancer drug comprising administering the drug to a transgenic animal expressing or over-expressing the prostate cancer modulatory protein, or an animal lacking the prostate cancer modulatory protein, for example as a result of a gene knockout.

Moreover, provided herein is a biochip comprising one or more nucleic acid segments of Tables 1-16, wherein the biochip comprises fewer than 1000 nucleic acid probes. Preferably, at least two nucleic acid segments are included. More preferably, at least three nucleic acid segments are included.

Furthermore, a method of diagnosing a disorder associated with prostate cancer is provided. The method comprises determining the expression of a gene of Tables 1-16, in a first tissue type of a first individual, and comparing the distribution to the expression of the gene from a second normal tissue type from the first individual or a second unaffected individual. A difference in the expression indicates that the first individual has a disorder associated with prostate cancer.

In a further embodiment, the biochip also includes a polynucleotide sequence of a gene that is not up- and down-regulated in prostate cancer.

In one embodiment a method for screening for a bioactive agent capable of interfering with the binding of a prostate cancer modulating protein (prostate cancer modulatory protein) or a fragment thereof and an antibody which binds to said prostate cancer modulatory protein or fragment thereof. In a preferred embodiment, the method comprises combining a prostate cancer modulatory protein or fragment thereof, a candidate bioactive agent and an antibody which binds to said prostate cancer modulatory protein or fragment thereof. The method further includes determining the binding of said prostate cancer modulatory protein or fragment thereof and said antibody. Wherein there is a change in binding, an agent is identified as an interfering agent. The interfering agent can be an agonist or an antagonist. Preferably, the agent inhibits prostate cancer.

Also provided herein are methods of eliciting an immune response in an individual. In one embodiment a method provided herein comprises administering to an

individual a composition comprising a prostate cancer modulating protein, or a fragment thereof. In another embodiment, the protein is encoded by a nucleic acid selected from those of Tables 1-16.

Further provided herein are compositions capable of eliciting an immune response in an individual. In one embodiment, a composition provided herein comprises a prostate cancer modulating protein, preferably encoded by a nucleic acid of Tables 1-16, or a fragment thereof, and a pharmaceutically acceptable carrier. In another embodiment, said composition comprises a nucleic acid comprising a sequence encoding a prostate cancer modulating protein, preferably selected from the nucleic acids of Tables 1-16, and a pharmaceutically acceptable carrier.

Also provided are methods of neutralizing the effect of a prostate cancer protein, or a fragment thereof, comprising contacting an agent specific for said protein with said protein in an amount sufficient to effect neutralization. In another embodiment, the protein is encoded by a nucleic acid selected from those of Tables 1-16.

In another aspect of the invention, a method of treating an individual for prostate cancer is provided. In one embodiment, the method comprises administering to said individual an inhibitor of a prostate cancer modulating protein. In another embodiment, the method comprises administering to a patient having prostate cancer an antibody to a prostate cancer modulating protein conjugated to a therapeutic moiety. Such a therapeutic moiety can be a cytotoxic agent or a radioisotope.

DETAILED DESCRIPTION OF THE INVENTION

In accordance with the objects outlined above, the present invention provides novel methods for diagnosis and prognosis evaluation for prostate cancer (PC), including metastatic prostate cancer, as well as methods for screening for compositions which modulate prostate cancer. Also provided are methods for treating prostate cancer.

In addition to the other nucleic acid and peptide sequences, the present invention also relates to the identification of PAA2 as a gene that is highly over expressed in prostate cancer patient tissues. PAA2 sequence is identical to the zinc transporter ZNT4. Results presented herein demonstrate that PAA2/ZNT4 is highly expressed in prostate cancer cells. The prostate gland is unique in that it has the highest capacity of any organ in the body

to accumulate zinc. Zinc uptake is regulated by prolactin and testosterone, which induce the expression of a member of the ZIP family of zinc transporters (Costello et al., 1999, J. Biol. Chem. 274:17499-17504). Zinc accumulation in the prostate functions to inhibit citrate oxidation, which results in a decrease in cellular ATP production (Costello and Franklin, 1998, Prostate 35:285-296). Cancer cells are more sensitive to decreased ATP production and have evolved to prevent zinc accumulation. Without wishing to be bound by theory, the up-regulation of ZNT4 in prostate cancer cells may result in protection of the cells from high zinc levels by its ability to pump accumulated zinc out of the cells.

The present invention also relates to nucleic acid sequences encoding PBH1.

PBH1 is related to human TRPC7 (transient receptor potential-related channels, NP_003298), a putative calcium channel highly expressed in brain (Nagamine et al., Genomics 54:124-131 (1998)). Trp is related to melastatin, a gene down-regulated in metastatic melanomas (Duncan et al., Cancer Res. 58:1515-1520 (1998)), and MTR1, a gene localized to within the Beckwith-Wiedemann syndrome/Wilm's tumor susceptibility region (Prawitt et al., Hum. Mol. Genet. 9:203-216 (2000)). Without wishing to be bound by theory, it is believed that PBH1 functions as a calcium channel.

As a calcium channel, PBH1 is an ideal target for a small molecule therapeutic, or a therapeutic antibody that disrupts channel function. CD20, the target of Rituximab in non-Hodgkin's lymphoma (Maloney et al., Blood 90:2188-2195 (1997); Leget and Czuczman, Curr. Opin. Oncol. 10:548-551 (1998)), is a plasma membrane calcium channel expressed in B cells (Tedder and Engel, Immunol. Today 15:450-454 (1994)). Similarly, a small molecule, or antibody that inhibits or alters a calcium signal mediated by PBH1, will result in the death of prostate cancer cells.

PBH1, and other genes of the invention, are also be useful as targets for cytotoxic T-lymphocytes. Genes that are tumor specific, or that are expressed in immune-privileged organs, are currently being used as potential vaccine targets (Van den Eynde and Boon, Int. J. Clin. Lab. Res. 27:81-86 (1997)). The expression pattern of PBH1 indicates that it is an ideal target for cytotoxic T-lymphocytes. Thus, therapies that utilize PBH1-specific cytotoxic T-lymphocytes to induce prostate cancer cell death are also provided by this invention. See, e.g., U.S. Patent No. 6,051,227 and WO 00/32231, the disclosures of which are herein incorporated by reference.

The present invention is also related to the identification of PAA3 as a gene that is important in the modulation of prostate cancer and or breast cancer.

Tables 1-16 provide unigene cluster identification numbers, exemplar accession numbers, or genomic nucleotide position numbers for the nucleotide sequence of genes that exhibit increased or decreased expression in prostate cancer samples.

Definitions

The term “prostate cancer protein” or “prostate cancer polynucleotide” or “prostate cancer-associated transcript” refers to nucleic acid and polypeptide polymorphic variants, alleles, mutants, and interspecies homologues that: (1) have a nucleotide sequence that has greater than about 60% nucleotide sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater nucleotide sequence identity, preferably over a region of over a region of at least about 25, 50, 100, 200, 500, 1000, or more nucleotides, to a nucleotide sequence of or associated with a unigene cluster of Tables 1-16; (2) bind to antibodies, e.g., polyclonal antibodies, raised against an immunogen comprising an amino acid sequence encoded by a nucleotide sequence of or associated with a unigene cluster of Tables 1-16, and conservatively modified variants thereof; (3) specifically hybridize under stringent hybridization conditions to a nucleic acid sequence, or the complement thereof of Tables 1-16 and conservatively modified variants thereof or (4) have an amino acid sequence that has greater than about 60% amino acid sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater amino sequence identity, preferably over a region of over a region of at least about 25, 50, 100, 200, 500, 1000, or more amino acid, to an amino acid sequence encoded by a nucleotide sequence of or associated with a unigene cluster of Tables 1-16. A polynucleotide or polypeptide sequence is typically from a mammal including, but not limited to, primate, e.g., human; rodent, e.g., rat, mouse, hamster; cow, pig, horse, sheep, or other mammal. A “prostate cancer polypeptide” and a “prostate cancer polynucleotide,” include both naturally occurring or recombinant forms.

A “full length” prostate cancer protein or nucleic acid refers to a prostate cancer polypeptide or polynucleotide sequence, or a variant thereof, that contains all of the elements normally contained in one or more naturally occurring, wild type prostate cancer

polynucleotide or polypeptide sequences. For example, a full length prostate cancer nucleic acid will typically comprise all of the exons that encode for the full length, naturally occurring protein. The “full length” may be prior to, or after, various stages of post-translation processing or splicing, including alternative splicing.

5 “Biological sample” as used herein is a sample of biological tissue or fluid that contains nucleic acids or polypeptides, e.g., of a prostate cancer protein, polynucleotide or transcript. Such samples include, but are not limited to, tissue isolated from primates, e.g., humans, or rodents, e.g., mice, and rats. Biological samples may also include sections of tissues such as biopsy and autopsy samples, frozen sections taken for histologic purposes,
10 blood, plasma, serum, sputum, stool, tears, mucus, hair, skin, etc. Biological samples also include explants and primary and/or transformed cell cultures derived from patient tissues. A biological sample is typically obtained from a eukaryotic organism, most preferably a mammal such as a primate e.g., chimpanzee or human; cow; dog; cat; a rodent, e.g., guinea pig, rat, mouse; rabbit; or a bird; reptile; or fish.

15 “Providing a biological sample” means to obtain a biological sample for use in methods described in this invention. Most often, this will be done by removing a sample of cells from an animal, but can also be accomplished by using previously isolated cells (e.g., isolated by another person, at another time, and/or for another purpose), or by performing the methods of the invention *in vivo*. Archival tissues, having treatment or outcome history, will
20 be particularly useful.

The terms “identical” or percent “identity,” in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same (i.e., about 60% identity, preferably 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%,
25 95%, 96%, 97%, 98%, 99%, or higher identity over a specified region, when compared and aligned for maximum correspondence over a comparison window or designated region) as measured using a BLAST or BLAST 2.0 sequence comparison algorithms with default parameters described below, or by manual alignment and visual inspection (*see, e.g.*, NCBI web site <http://www.ncbi.nlm.nih.gov/BLAST/> or the like). Such sequences are then said to
30 be “substantially identical.” This definition also refers to, or may be applied to, the compliment of a test sequence. The definition also includes sequences that have deletions

and/or additions, as well as those that have substitutions, as well as naturally occurring, e.g., polymorphic or allelic variants, and man-made variants. As described below, the preferred algorithms can account for gaps and the like. Preferably, identity exists over a region that is at least about 25 amino acids or nucleotides in length, or more preferably over a region that is
5 50-100 amino acids or nucleotides in length.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Preferably, default
10 program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

A “comparison window”, as used herein, includes reference to a segment of one of the number of contiguous positions selected from the group consisting typically of
15 from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat’l. Acad. Sci. USA* 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and
25 visual inspection (*see, e.g., Current Protocols in Molecular Biology* (Ausubel *et al.*, eds. 1995 supplement)).

Preferred examples of algorithms that are suitable for determining percent sequence identity and sequence similarity include the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al.*, *Nuc. Acids Res.* 25:3389-3402 (1977) and Altschul *et al.*, *J. Mol. Biol.* 215:403-410 (1990). BLAST and BLAST 2.0 are used, with the parameters
30 described herein, to determine percent sequence identity for the nucleic acids and proteins of

the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, e.g., for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W , T , and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, $M=5$, $N=-4$ and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)) alignments (B) of 50, expectation (E) of 10, $M=5$, $N=-4$, and a comparison of both strands.

The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin & Altschul, *Proc. Nat'l. Acad. Sci. USA* 90:5873-5787 (1993)). One measure of similarity provided by the BLAST algorithm is the smallest sum probability ($P(N)$), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, more preferably less than about 0.01, and most preferably less than about 0.001. Log values may be large negative numbers, e.g., 5, 10, 20, 30, 40, 40, 70, 90, 110, 150, 170, etc.

An indication that two nucleic acid sequences or polypeptides are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the antibodies raised against the polypeptide encoded by the second nucleic acid, as described below. Thus, a polypeptide is typically substantially identical to a second polypeptide, e.g., where the two peptides differ only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules or their complements hybridize to each other under stringent conditions, as described below. Yet another indication that two nucleic acid sequences are substantially identical is that the same primers can be used to amplify the sequences.

A "host cell" is a naturally occurring cell or a transformed cell that contains an expression vector and supports the replication or expression of the expression vector. Host cells may be cultured cells, explants, cells *in vivo*, and the like. Host cells may be prokaryotic cells such as *E. coli*, or eukaryotic cells such as yeast, insect, amphibian, or mammalian cells such as CHO, HeLa, and the like (*see, e.g.*, the American Type Culture Collection catalog or web site, www.atcc.org).

The terms "isolated," "purified," or "biologically pure" refer to material that is substantially or essentially free from components that normally accompany it as found in its native state. Purity and homogeneity are typically determined using analytical chemistry techniques such as polyacrylamide gel electrophoresis or high performance liquid chromatography. A protein or nucleic acid that is the predominant species present in a preparation is substantially purified. In particular, an isolated nucleic acid is separated from some open reading frames that naturally flank the gene and encode proteins other than protein encoded by the gene. The term "purified" in some embodiments denotes that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. Preferably, it means that the nucleic acid or protein is at least 85% pure, more preferably at least 95% pure, and most preferably at least 99% pure. "Purify" or "purification" in other embodiments means removing at least one contaminant from the composition to be purified. In this sense, purification does not require that the purified compound be homogenous, e.g., 100% pure.

The terms "polypeptide," "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding

naturally occurring amino acid, as well as to naturally occurring amino acid polymers, those containing modified residues, and non-naturally occurring amino acid polymer.

The term “amino acid” refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function similarly to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline, γ -carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, e.g., an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs may have modified R groups (e.g., norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refers to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that functions similarly to a naturally occurring amino acid.

Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

“Conservatively modified variants” applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical or associated, e.g., naturally contiguous, sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode most proteins. For instance, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to another of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are “silent variations,” which are one species of conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes silent variations of the nucleic acid. One of skill will recognize that in certain contexts each codon in a nucleic acid (except AUG, which is ordinarily the

only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule. Accordingly, often silent variations of a nucleic acid which encodes a polypeptide is implicit in a described sequence with respect to the expression product, but not with respect to actual probe sequences.

5 As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid.

10 Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention. typically conservative substitutions for one another: 1) Alanine (A), Glycine (G); 2) Aspartic acid (D), Glutamic acid (E); 3) Asparagine (N), Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W); 7) Serine (S), Threonine (T); and 8) Cysteine (C), Methionine (M) (*see, e.g., Creighton, Proteins* (1984)).

20 Macromolecular structures such as polypeptide structures can be described in terms of various levels of organization. For a general discussion of this organization, *see, e.g., Alberts et al., Molecular Biology of the Cell* (3rd ed., 1994) and Cantor & Schimmel, *Biophysical Chemistry Part I: The Conformation of Biological Macromolecules* (1980). "Primary structure" refers to the amino acid sequence of a particular peptide. "Secondary structure" refers to locally ordered, three dimensional structures within a polypeptide. These structures are commonly known as domains. Domains are portions of a polypeptide that often form a compact unit of the polypeptide and are typically 25 to approximately 500 amino acids long. Typical domains are made up of sections of lesser organization such as stretches of β -sheet and α -helices. "Tertiary structure" refers to the complete three dimensional structure of a polypeptide monomer. "Quaternary structure" refers to the three dimensional structure formed, usually by the noncovalent association of independent tertiary units. Anisotropic terms are also known as energy terms.

30

“Nucleic acid” or “oligonucleotide” or “polynucleotide” or grammatical equivalents used herein means at least two nucleotides covalently linked together.

Oligonucleotides are typically from about 5, 6, 7, 8, 9, 10, 12, 15, 25, 30, 40, 50 or more nucleotides in length, up to about 100 nucleotides in length. Nucleic acids and

5 polynucleotides are a polymers of any length, including longer lengths, e.g., 200, 300, 500, 1000, 2000, 3000, 5000, 7000, 10,000, etc. A nucleic acid of the present invention will generally contain phosphodiester bonds, although in some cases, nucleic acid analogs are included that may have alternate backbones, comprising, e.g., phosphoramidate,

phosphorothioate, phosphorodithioate, or O-methylphosphoroamidite linkages (see Eckstein,

10 Oligonucleotides and Analogues: A Practical Approach, Oxford University Press); and peptide nucleic acid backbones and linkages. Other analog nucleic acids include those with

positive backbones; non-ionic backbones, and non-ribose backbones, including those described in U.S. Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7, ASC

Symposium Series 580, *Carbohydrate Modifications in Antisense Research*, Sanghui &

15 Cook, eds.. Nucleic acids containing one or more carbocyclic sugars are also included within one definition of nucleic acids. Modifications of the ribose-phosphate backbone may be done

for a variety of reasons, e.g. to increase the stability and half-life of such molecules in physiological environments or as probes on a biochip. Mixtures of naturally occurring

nucleic acids and analogs can be made; alternatively, mixtures of different nucleic acid

20 analogs, and mixtures of naturally occurring nucleic acids and analogs may be made.

A variety of references disclose such nucleic acid analogs, including, for example, phosphoramidate (Beaucage et al., Tetrahedron 49(10):1925 (1993) and references therein; Letsinger, J. Org. Chem. 35:3800 (1970); Sprinzl et al., Eur. J. Biochem. 81:579

(1977); Letsinger et al., Nucl. Acids Res. 14:3487 (1986); Sawai et al, Chem. Lett. 805

25 (1984), Letsinger et al., J. Am. Chem. Soc. 110:4470 (1988); and Pauwels et al., *Chemica Scripta* 26:141 (1986)), phosphorothioate (Mag et al., Nucleic Acids Res. 19:1437 (1991);

and U.S. Patent No. 5,644,048), phosphorodithioate (Briu et al., J. Am. Chem. Soc. 111:2321

(1989), O-methylphosphoroamidite linkages (see Eckstein, Oligonucleotides and Analogues: A Practical Approach, Oxford University Press), and peptide nucleic acid backbones and

30 linkages (see Egholm, J. Am. Chem. Soc. 114:1895 (1992); Meier et al., Chem. Int. Ed. Engl. 31:1008 (1992); Nielsen, Nature, 365:566 (1993); Carlsson et al., Nature 380:207 (1996), all

of which are incorporated by reference). Other analog nucleic acids include those with positive backbones (Denpcy et al., Proc. Natl. Acad. Sci. USA 92:6097 (1995); non-ionic backbones (U.S. Patent Nos. 5,386,023, 5,637,684, 5,602,240, 5,216,141 and 4,469,863; Kiedrowshi et al., Angew. Chem. Intl. Ed. English 30:423 (1991); Letsinger et al., J. Am. Chem. Soc. 110:4470 (1988); Letsinger et al., Nucleoside & Nucleotide 13:1597 (1994); Chapters 2 and 3, ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research", Ed. Y.S. Sanghui and P. Dan Cook; Mesmaeker et al., Bioorganic & Medicinal Chem. Lett. 4:395 (1994); Jeffs et al., J. Biomolecular NMR 34:17 (1994); Tetrahedron Lett. 37:743 (1996)) and non-ribose backbones, including those described in U.S. Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7, ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research", Ed. Y.S. Sanghui and P. Dan Cook. Nucleic acids containing one or more carbocyclic sugars are also included within one definition of nucleic acids (see Jenkins et al., Chem. Soc. Rev. (1995) pp 169-176). Several nucleic acid analogs are described in Rawls, C & E News June 2, 1997 page 35. All of these references are hereby expressly incorporated by reference.

Particularly preferred are peptide nucleic acids (PNA) which includes peptide nucleic acid analogs. These backbones are substantially non-ionic under neutral conditions, in contrast to the highly charged phosphodiester backbone of naturally occurring nucleic acids. This results in two advantages. First, the PNA backbone exhibits improved hybridization kinetics. PNAs have larger changes in the melting temperature (T_m) for mismatched versus perfectly matched basepairs. DNA and RNA typically exhibit a 2-4°C drop in T_m for an internal mismatch. With the non-ionic PNA backbone, the drop is closer to 7-9°C. Similarly, due to their non-ionic nature, hybridization of the bases attached to these backbones is relatively insensitive to salt concentration. In addition, PNAs are not degraded by cellular enzymes, and thus can be more stable.

The nucleic acids may be single stranded or double stranded, as specified, or contain portions of both double stranded or single stranded sequence. As will be appreciated by those in the art, the depiction of a single strand also defines the sequence of the complementary strand; thus the sequences described herein also provide the complement of the sequence. The nucleic acid may be DNA, both genomic and cDNA, RNA or a hybrid, where the nucleic acid may contain combinations of deoxyribo- and ribo-nucleotides, and

combinations of bases, including uracil, adenine, thymine, cytosine, guanine, inosine, xanthine hypoxanthine, isocytosine, isoguanine, etc. "Transcript" typically refers to a naturally occurring RNA, e.g., a pre-mRNA, hnRNA, or mRNA. As used herein, the term "nucleoside" includes nucleotides and nucleoside and nucleotide analogs, and modified
5 nucleosides such as amino modified nucleosides. In addition, "nucleoside" includes non-naturally occurring analog structures. Thus, e.g. the individual units of a peptide nucleic acid, each containing a base, are referred to herein as a nucleoside.

A "label" or a "detectable moiety" is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, chemical, or other physical
10 means. For example, useful labels include fluorescent dyes, electron-dense reagents, enzymes (e.g., as commonly used in an ELISA), biotin, digoxigenin, or haptens and proteins or other entities which can be made detectable, e.g., by incorporating a radiolabel into the peptide or used to detect antibodies specifically reactive with the peptide. The radioisotope may be, for example, ^3H , ^{14}C , ^{32}P , ^{35}S , or ^{125}I . In some cases, particularly using antibodies against the
15 proteins of the invention, the radioisotopes are used as toxic moieties, as described below. The labels may be incorporated into the prostate cancer nucleic acids, proteins and antibodies at any position. Any method known in the art for conjugating the antibody to the label may be employed, including those methods described by Hunter et al., Nature, 144:945 (1962); David et al., Biochemistry, 13:1014 (1974); Pain et al., J. Immunol. Meth., 40:219 (1981);
20 and Nygren, J. Histochem. and Cytochem., 30:407 (1982). The lifetime of radiolabeled peptides or radiolabeled antibody compositions may extended by the addition of substances that stabilize the radiolabeled peptide or antibody and protect it from degradation. Any substance or combination of substances that stabilize the radiolabeled peptide or antibody may be used including those substances disclosed in US Patent No. 5,961,955.

25 An "effector" or "effector moiety" or "effector component" is a molecule that is bound (or linked, or conjugated), either covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds, to an antibody. The "effector" can be a variety of molecules including, e.g., detection moieties including radioactive compounds, fluorescent compounds, an enzyme or substrate, tags such as epitope
30 tags, a toxin; activatable moieties, a chemotherapeutic agent; a lipase; an antibiotic; or a radioisotope emitting "hard" e.g., beta radiation.

A "labeled nucleic acid probe or oligonucleotide" is one that is bound, either covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds to a label such that the presence of the probe may be detected by detecting the presence of the label bound to the probe. Alternatively, method
5 using high affinity interactions may achieve the same results where one of a pair of binding partners binds to the other, e.g., biotin, streptavidin.

As used herein a "nucleic acid probe or oligonucleotide" is defined as a nucleic acid capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, usually
10 through hydrogen bond formation. As used herein, a probe may include natural (i.e., A, G, C, or T) or modified bases (7-deazaguanosine, inosine, etc.). In addition, the bases in a probe may be joined by a linkage other than a phosphodiester bond, so long as it does not functionally interfere with hybridization. Thus, e.g., probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages.
15 It will be understood by one of skill in the art that probes may bind target sequences lacking complete complementarity with the probe sequence depending upon the stringency of the hybridization conditions. The probes are preferably directly labeled as with isotopes, chromophores, lumiphores, chromogens, or indirectly labeled such as with biotin to which a streptavidin complex may later bind. By assaying for the presence or absence of the probe,
20 one can detect the presence or absence of the select sequence or subsequence. Diagnosis or prognosis may be based at the genomic level, or at the level of RNA or protein expression.

The term "recombinant" when used with reference, e.g., to a cell, or nucleic acid, protein, or vector, indicates that the cell, nucleic acid, protein or vector, has been modified by the introduction of a heterologous nucleic acid or protein or the alteration of a
25 native nucleic acid or protein, or that the cell is derived from a cell so modified. Thus, e.g., recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all. By the term "recombinant nucleic acid" herein is meant nucleic acid, originally formed *in vitro*, in general, by the manipulation of nucleic acid, e.g., using
30 polymerases and endonucleases, in a form not normally found in nature. In this manner, operably linkage of different sequences is achieved. Thus an isolated nucleic acid, in a linear

form, or an expression vector formed *in vitro* by ligating DNA molecules that are not normally joined, are both considered recombinant for the purposes of this invention. It is understood that once a recombinant nucleic acid is made and reintroduced into a host cell or organism, it will replicate non-recombinantly, i.e., using the *in vivo* cellular machinery of the host cell rather than *in vitro* manipulations; however, such nucleic acids, once produced recombinantly, although subsequently replicated non-recombinantly, are still considered recombinant for the purposes of the invention. Similarly, a “recombinant protein” is a protein made using recombinant techniques, i.e., through the expression of a recombinant nucleic acid as depicted above.

The term “heterologous” when used with reference to portions of a nucleic acid indicates that the nucleic acid comprises two or more subsequences that are not normally found in the same relationship to each other in nature. For instance, the nucleic acid is typically recombinantly produced, having two or more sequences, e.g., from unrelated genes arranged to make a new functional nucleic acid, e.g., a promoter from one source and a coding region from another source. Similarly, a heterologous protein will often refer to two or more subsequences that are not found in the same relationship to each other in nature (e.g., a fusion protein).

A “promoter” is defined as an array of nucleic acid control sequences that direct transcription of a nucleic acid. As used herein, a promoter includes necessary nucleic acid sequences near the start site of transcription, such as, in the case of a polymerase II type promoter, a TATA element. A promoter also optionally includes distal enhancer or repressor elements, which can be located as much as several thousand base pairs from the start site of transcription. A “constitutive” promoter is a promoter that is active under most environmental and developmental conditions. An “inducible” promoter is a promoter that is active under environmental or developmental regulation. The term “operably linked” refers to a functional linkage between a nucleic acid expression control sequence (such as a promoter, or array of transcription factor binding sites) and a second nucleic acid sequence, wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence.

An “expression vector” is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements that permit transcription of a

particular nucleic acid in a host cell. The expression vector can be part of a plasmid, virus, or nucleic acid fragment. Typically, the expression vector includes a nucleic acid to be transcribed operably linked to a promoter.

The phrase “selectively (or specifically) hybridizes to” refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence that is determinative of the presence of the nucleotide sequence, in a heterogeneous population of nucleic acids and other biologics (e.g., total cellular or library DNA or RNA). Similarly, the phrase “specifically (or selectively) binds” to an antibody or “specifically (or selectively) immunoreactive with,” when referring to a protein or peptide, refers to a binding reaction that is determinative of the presence of the protein, in a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay or nucleic acid hybridization conditions, the specified antibodies or nucleic acid probes bind to a particular protein nucleotide sequences at least two times the background and more typically more than 10 to 100 times background.

Specific binding to an antibody under such conditions requires an antibody that is selected for its specificity for a particular protein. For example, polyclonal antibodies raised to a particular protein, polymorphic variants, alleles, orthologs, and conservatively modified variants, or splice variants, or portions thereof, can be selected to obtain only those polyclonal antibodies that are specifically immunoreactive with the desired prostate cancer protein and not with other proteins. This selection may be achieved by subtracting out antibodies that cross-react with other molecules. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays are routinely used to select antibodies specifically immunoreactive with a protein (*see, e.g., Harlow & Lane, Antibodies, A Laboratory Manual* (1988) for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity).

The phrase “stringent hybridization conditions” refers to conditions under which a probe will hybridize to its target subsequence, typically in a complex mixture of nucleic acids, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in

Tijssen, *Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes*, "Overview of principles of hybridization and the strategy of nucleic acid assays"

(1993). Generally, stringent conditions are selected to be about 5-10°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength pH. The T_m is

5 the temperature (under defined ionic strength, pH, and nucleic concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at T_m , 50% of the probes are occupied at equilibrium). Stringent conditions will be those in which the salt concentration is less than

10 about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g., 10 to 50 nucleotides) and at least about 60°C for long probes (e.g., greater than 50 nucleotides).

Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. For selective or specific hybridization, a positive signal is at least two times background, preferably 10 times background hybridization. Exemplary stringent

15 hybridization conditions can be as following: 50% formamide, 5x SSC, and 1% SDS, incubating at 42°C, or, 5x SSC, 1% SDS, incubating at 65°C, with wash in 0.2x SSC, and 0.1% SDS at 65°C. For PCR, a temperature of about 36°C is typical for low stringency amplification, although annealing temperatures may vary between about 32°C and 48°C

depending on primer length. For high stringency PCR amplification, a temperature of about

20 62°C is typical, although high stringency annealing temperatures can range from about 50°C to about 65°C, depending on the primer length and specificity. Typical cycle conditions for both high and low stringency amplifications include a denaturation phase of 90°C - 95°C for

30 sec - 2 min., an annealing phase lasting 30 sec. - 2 min., and an extension phase of about 72°C for 1 - 2 min. Protocols and guidelines for low and high stringency amplification

reactions are provided, e.g., in Innis *et al.* (1990) *PCR Protocols, A Guide to Methods and*

25 *Applications*, Academic Press, Inc. N.Y.).

Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical.

This occurs, e.g., when a copy of a nucleic acid is created using the maximum codon

30 degeneracy permitted by the genetic code. In such cases, the nucleic acids typically hybridize

under moderately stringent hybridization conditions. Exemplary “moderately stringent hybridization conditions” include a hybridization in a buffer of 40% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 1X SSC at 45°C. A positive hybridization is at least twice background. Those of ordinary skill will readily recognize that alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency. Additional guidelines for determining hybridization parameters are provided in numerous reference, e.g., and Current Protocols in Molecular Biology, ed. Ausubel, *et al.*

The phrase “functional effects” in the context of assays for testing compounds that modulate activity of a prostate cancer protein includes the determination of a parameter that is indirectly or directly under the influence of the prostate cancer protein or nucleic acid, e.g., a functional, physical, or chemical effect, such as the ability to decrease prostate cancer. It includes ligand binding activity; cell growth on soft agar; anchorage dependence; contact inhibition and density limitation of growth; cellular proliferation; cellular transformation; growth factor or serum dependence; tumor specific marker levels; invasiveness into Matrigel; tumor growth and metastasis *in vivo*; mRNA and protein expression in cells undergoing metastasis, and other characteristics of prostate cancer cells. “Functional effects” include *in vitro*, *in vivo*, and *ex vivo* activities.

By “determining the functional effect” is meant assaying for a compound that increases or decreases a parameter that is indirectly or directly under the influence of a prostate cancer protein sequence, e.g., functional, enzymatic, physical and chemical effects. Such functional effects can be measured by any means known to those skilled in the art, e.g., changes in spectroscopic characteristics (e.g., fluorescence, absorbance, refractive index), hydrodynamic (e.g., shape), chromatographic, or solubility properties for the protein, measuring inducible markers or transcriptional activation of the prostate cancer protein; measuring binding activity or binding assays, e.g. binding to antibodies or other ligands, and measuring cellular proliferation. Determination of the functional effect of a compound on prostate cancer can also be performed using prostate cancer assays known to those of skill in the art such as an *in vitro* assays, e.g., cell growth on soft agar; anchorage dependence; contact inhibition and density limitation of growth; cellular proliferation; cellular transformation; growth factor or serum dependence; tumor specific marker levels; invasiveness into Matrigel; tumor growth and metastasis *in vivo*; mRNA and protein

expression in cells undergoing metastasis, and other characteristics of prostate cancer cells. The functional effects can be evaluated by many means known to those skilled in the art, e.g., microscopy for quantitative or qualitative measures of alterations in morphological features, measurement of changes in RNA or protein levels for prostate cancer-associated sequences, measurement of RNA stability, identification of downstream or reporter gene expression (CAT, luciferase, β -gal, GFP and the like), e.g., via chemiluminescence, fluorescence, colorimetric reactions, antibody binding, inducible markers, and ligand binding assays.

“Inhibitors”, “activators”, and “modulators” of prostate cancer polynucleotide and polypeptide sequences are used to refer to activating, inhibitory, or modulating molecules or compounds identified using *in vitro* and *in vivo* assays of prostate cancer polynucleotide and polypeptide sequences. Inhibitors are compounds that, e.g., bind to, partially or totally block activity, decrease, prevent, delay activation, inactivate, desensitize, or down regulate the activity or expression of prostate cancer proteins, e.g., antagonists. Antisense nucleic acids may seem to inhibit expression and subsequent function of the protein. “Activators” are compounds that increase, open, activate, facilitate, enhance activation, sensitize, agonize, or up regulate prostate cancer protein activity. Inhibitors, activators, or modulators also include genetically modified versions of prostate cancer proteins, e.g., versions with altered activity, as well as naturally occurring and synthetic ligands, antagonists, agonists, antibodies, small chemical molecules and the like. Such assays for inhibitors and activators include, e.g., expressing the prostate cancer protein *in vitro*, in cells, or cell membranes, applying putative modulator compounds, and then determining the functional effects on activity, as described above. Activators and inhibitors of prostate cancer can also be identified by incubating prostate cancer cells with the test compound and determining increases or decreases in the expression of 1 or more prostate cancer proteins, e.g., 1; 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50 or more prostate cancer proteins, such as prostate cancer proteins encoded by the sequences set out in Tables 1-16.

Samples or assays comprising prostate cancer proteins that are treated with a potential activator, inhibitor, or modulator are compared to control samples without the inhibitor, activator, or modulator to examine the extent of inhibition. Control samples (untreated with inhibitors) are assigned a relative protein activity value of 100%. Inhibition of a polypeptide is achieved when the activity value relative to the control is about 80%,

preferably 50%, more preferably 25-50%. Activation of a prostate cancer polypeptide is achieved when the activity value relative to the control (untreated with activators) is 110%, more preferably 150%, more preferably 200-500% (i.e., two to five fold higher relative to the control), more preferably 1000-3000% higher.

5 The phrase "changes in cell growth" refers to any change in cell growth and proliferation characteristics *in vitro* or *in vivo*, such as formation of foci, anchorage independence, semi-solid or soft agar growth, changes in contact inhibition and density limitation of growth, loss of growth factor or serum requirements, changes in cell morphology, gaining or losing immortalization, gaining or losing tumor specific markers, 10 ability to form or suppress tumors when injected into suitable animal hosts, and/or immortalization of the cell. *See, e.g., Freshney, Culture of Animal Cells a Manual of Basic Technique* pp. 231-241 (3rd ed. 1994).

 "Tumor cell" refers to precancerous, cancerous, and normal cells in a tumor.

 "Cancer cells," "transformed" cells or "transformation" in tissue culture, refers 15 to spontaneous or induced phenotypic changes that do not necessarily involve the uptake of new genetic material. Although transformation can arise from infection with a transforming virus and incorporation of new genomic DNA, or uptake of exogenous DNA, it can also arise spontaneously or following exposure to a carcinogen, thereby mutating an endogenous gene. Transformation is associated with phenotypic changes, such as immortalization of cells, 20 aberrant growth control, nonmorphological changes, and/or malignancy (*see, Freshney, Culture of Animal Cells a Manual of Basic Technique* (3rd ed. 1994)).

 "Antibody" refers to a polypeptide comprising a framework region from an immunoglobulin gene or fragments thereof that specifically binds and recognizes an antigen. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, 25 epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively. Typically, the antigen-binding region of an antibody or its functional equivalent will be most critical in specificity and affinity of binding. *See Paul,* 30 *Fundamental Immunology.*

An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one "light" (about 25 kD) and one "heavy" chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible for antigen recognition. The terms variable light chain (V_L) and variable heavy chain (V_H) refer to these light and heavy chains respectively.

Antibodies exist, e.g., as intact immunoglobulins or as a number of well-characterized fragments produced by digestion with various peptidases. Thus, e.g., pepsin digests an antibody below the disulfide linkages in the hinge region to produce $F(ab)'_2$, a dimer of Fab which itself is a light chain joined to V_H - C_H1 by a disulfide bond. The $F(ab)'_2$ may be reduced under mild conditions to break the disulfide linkage in the hinge region, thereby converting the $F(ab)'_2$ dimer into an Fab' monomer. The Fab' monomer is essentially Fab with part of the hinge region (*see Fundamental Immunology* (Paul ed., 3d ed. 1993)). While various antibody fragments are defined in terms of the digestion of an intact antibody, one of skill will appreciate that such fragments may be synthesized *de novo* either chemically or by using recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments either produced by the modification of whole antibodies, or those synthesized *de novo* using recombinant DNA methodologies (e.g., single chain Fv) or those identified using phage display libraries (*see, e.g., McCafferty et al., Nature* 348:552-554 (1990)).

For preparation of antibodies, e.g., recombinant, monoclonal, or polyclonal antibodies, many technique known in the art can be used (*see, e.g., Kohler & Milstein, Nature* 256:495-497 (1975); Kozbor *et al., Immunology Today* 4:72 (1983); Cole *et al.*, pp. 77-96 in *Monoclonal Antibodies and Cancer Therapy* (1985); Coligan, *Current Protocols in Immunology* (1991); Harlow & Lane, *Antibodies, A Laboratory Manual* (1988); and Goding, *Monoclonal Antibodies: Principles and Practice* (2d ed. 1986)). Techniques for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies. Alternatively, phage display technology can be used to identify antibodies and heteromeric Fab fragments that

specifically bind to selected antigens (*see, e.g.,* McCafferty *et al.*, *Nature* 348:552-554 (1990); Marks *et al.*, *Biotechnology* 10:779-783 (1992)).

A “chimeric antibody” is an antibody molecule in which (a) the constant region, or a portion thereof, is altered, replaced or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, *e.g.*, an enzyme, toxin, hormone, growth factor, drug, etc.; or (b) the variable region, or a portion thereof, is altered, replaced or exchanged with a variable region having a different or altered antigen specificity.

Identification of prostate cancer-associated sequences

In one aspect, the expression levels of genes are determined in different patient samples for which diagnosis information is desired, to provide expression profiles. An expression profile of a particular sample is essentially a “fingerprint” of the state of the sample; while two states may have any particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is characteristic of the state of the cell. That is, normal tissue (*e.g.*, normal prostate or other tissue) may be distinguished from cancerous or metastatic cancerous tissue of the prostate, or prostate cancer tissue or metastatic prostate cancerous tissue can be compared with tissue samples of prostate and other tissues from surviving cancer patients. By comparing expression profiles of tissue in known different prostate cancer states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained.

The identification of sequences that are differentially expressed in prostate cancer versus non-prostate cancer tissue allows the use of this information in a number of ways. For example, a particular treatment regime may be evaluated: does a chemotherapeutic drug act to down-regulate prostate cancer, and thus tumor growth or recurrence, in a particular patient. Similarly, diagnosis and treatment outcomes may be done or confirmed by comparing patient samples with the known expression profiles. Metastatic tissue can also be analyzed to determine the stage of prostate cancer in the tissue. Furthermore, these gene expression profiles (or individual genes) allow screening of drug candidates with an eye to

mimicking or altering a particular expression profile; e.g., screening can be done for drugs that suppress the prostate cancer expression profile. This may be done by making biochips comprising sets of the important prostate cancer genes, which can then be used in these screens. These methods can also be done on the protein basis; that is, protein expression levels of the prostate cancer proteins can be evaluated for diagnostic purposes or to screen candidate agents. In addition, the prostate cancer nucleic acid sequences can be administered for gene therapy purposes, including the administration of antisense nucleic acids, or the prostate cancer proteins (including antibodies and other modulators thereof) administered as therapeutic drugs.

Thus the present invention provides nucleic acid and protein sequences that are differentially expressed in prostate cancer, herein termed "prostate cancer sequences." As outlined below, prostate cancer sequences include those that are up-regulated (i.e., expressed at a higher level) in prostate cancer, as well as those that are down-regulated (i.e., expressed at a lower level). In a preferred embodiment, the prostate cancer sequences are from humans; however, as will be appreciated by those in the art, prostate cancer sequences from other organisms may be useful in animal models of disease and drug evaluation; thus, other prostate cancer sequences are provided, from vertebrates, including mammals, including rodents (rats, mice, hamsters, guinea pigs, etc.), primates, farm animals (including sheep, goats, pigs, cows, horses, etc.) and pets, e.g., (dogs, cats, etc.). Prostate cancer sequences from other organisms may be obtained using the techniques outlined below.

Prostate cancer sequences can include both nucleic acid and amino acid sequences. As will be appreciated by those in the art and is more fully outlined below, prostate cancer nucleic acid sequences are useful in a variety of applications, including diagnostic applications, which will detect naturally occurring nucleic acids, as well as screening applications; e.g., biochips comprising nucleic acid probes or PCR microtiter plates with selected probes to the prostate cancer sequences can be generated.

A prostate cancer sequence can be initially identified by substantial nucleic acid and/or amino acid sequence homology to the prostate cancer sequences outlined herein. Such homology can be based upon the overall nucleic acid or amino acid sequence, and is generally determined as outlined below, using either homology programs or hybridization conditions.

For identifying prostate cancer-associated sequences, the prostate cancer screen typically includes comparing genes identified in different tissues, e.g., normal and cancerous tissues, or tumor tissue samples from patients who have metastatic disease vs. non metastatic tissue. Other suitable tissue comparisons include comparing prostate cancer
5 samples with metastatic cancer samples from other cancers, such as lung, breast, gastrointestinal cancers, ovarian, etc. Samples of different stages of prostate cancer, e.g., survivor tissue, drug resistant states, and tissue undergoing metastasis, are applied to biochips comprising nucleic acid probes. The samples are first microdissected, if applicable, and treated as is known in the art for the preparation of mRNA. Suitable biochips are
10 commercially available, e.g. from Affymetrix. Gene expression profiles as described herein are generated and the data analyzed.

In one embodiment, the genes showing changes in expression as between normal and disease states are compared to genes expressed in other normal tissues, preferably normal prostate, but also including, and not limited to lung, heart, brain, liver, breast, kidney,
15 muscle, colon, small intestine, large intestine, spleen, bone and placenta. In a preferred embodiment, those genes identified during the prostate cancer screen that are expressed in any significant amount in other tissues are removed from the profile, although in some embodiments, this is not necessary. That is, when screening for drugs, it is usually preferable that the target be disease specific, to minimize possible side effects.

In a preferred embodiment, prostate cancer sequences are those that are up-regulated in prostate cancer; that is, the expression of these genes is higher in the prostate cancer tissue as compared to non-cancerous tissue. "Up-regulation" as used herein often means at least about a two-fold change, preferably at least about a three fold change, with at least about five-fold or higher being preferred. All unigene cluster identification numbers
20 and accession numbers herein are for the GenBank sequence database and the sequences of the accession numbers are hereby expressly incorporated by reference. GenBank is known in the art, *see, e.g.*, Benson, DA, *et al.*, Nucleic Acids Research 26:1-7 (1998) and <http://www.ncbi.nlm.nih.gov/>. Sequences are also available in other databases, e.g., European Molecular Biology Laboratory (EMBL) and DNA Database of Japan (DDBJ).

In another preferred embodiment, prostate cancer sequences are those that are down-regulated in prostate cancer; that is, the expression of these genes is lower in prostate
30

cancer tissue as compared to non-cancerous tissue (*see, e.g.*, Tables 8, 12 and 14). “Down-regulation” as used herein often means at least about a 1.5-fold change more preferably a two-fold change, preferably at least about a three fold change, with at least about five-fold or higher being most preferred.

5

Informatics

The ability to identify genes that are over or under expressed in prostate cancer can additionally provide high-resolution, high-sensitivity datasets which can be used in the areas of diagnostics, therapeutics, drug development, pharmacogenetics, protein structure, biosensor development, and other related areas. For example, the expression profiles can be used in diagnostic or prognostic evaluation of patients with prostate cancer. Or as another example, subcellular toxicological information can be generated to better direct drug structure and activity correlation (*see* Anderson, *Pharmaceutical Proteomics: Targets, Mechanism, and Function*, paper presented at the IBC Proteomics conference, Coronado, CA (June 11-12, 1998)). Subcellular toxicological information can also be utilized in a biological sensor device to predict the likely toxicological effect of chemical exposures and likely tolerable exposure thresholds (*see* U.S. Patent No. 5,811,231). Similar advantages accrue from datasets relevant to other biomolecules and bioactive agents (e.g., nucleic acids, saccharides, lipids, drugs, and the like).

Thus, in another embodiment, the present invention provides a database that includes at least one set of assay data. The data contained in the database is acquired, e.g., using array analysis either singly or in a library format. The database can be in substantially any form in which data can be maintained and transmitted, but is preferably an electronic database. The electronic database of the invention can be maintained on any electronic device allowing for the storage of and access to the database, such as a personal computer, but is preferably distributed on a wide area network, such as the World Wide Web.

The focus of the present section on databases that include peptide sequence data is for clarity of illustration only. It will be apparent to those of skill in the art that similar databases can be assembled for any assay data acquired using an assay of the invention.

The compositions and methods for identifying and/or quantitating the relative and/or absolute abundance of a variety of molecular and macromolecular species from a biological sample undergoing prostate cancer, i.e., the identification of prostate cancer-associated sequences described herein, provide an abundance of information, which can be correlated with pathological conditions, predisposition to disease, drug testing, therapeutic monitoring, gene-disease causal linkages, identification of correlates of immunity and physiological status, among others. Although the data generated from the assays of the invention is suited for manual review and analysis, in a preferred embodiment, prior data processing using high-speed computers is utilized.

An array of methods for indexing and retrieving biomolecular information is known in the art. For example, U.S. Patents 6,023,659 and 5,966,712 disclose a relational database system for storing biomolecular sequence information in a manner that allows sequences to be catalogued and searched according to one or more protein function hierarchies. U.S. Patent 5,953,727 discloses a relational database having sequence records containing information in a format that allows a collection of partial-length DNA sequences to be catalogued and searched according to association with one or more sequencing projects for obtaining full-length sequences from the collection of partial length sequences. U.S. Patent 5,706,498 discloses a gene database retrieval system for making a retrieval of a gene sequence similar to a sequence data item in a gene database based on the degree of similarity between a key sequence and a target sequence. U.S. Patent 5,538,897 discloses a method using mass spectroscopy fragmentation patterns of peptides to identify amino acid sequences in computer databases by comparison of predicted mass spectra with experimentally-derived mass spectra using a closeness-of-fit measure. U.S. Patent 5,926,818 discloses a multi-dimensional database comprising a functionality for multi-dimensional data analysis described as on-line analytical processing (OLAP), which entails the consolidation of projected and actual data according to more than one consolidation path or dimension. U.S. Patent 5,295,261 reports a hybrid database structure in which the fields of each database record are divided into two classes, navigational and informational data, with navigational fields stored in a hierarchical topological map which can be viewed as a tree structure or as the merger of two or more such tree structures.

See also Mount *et al.*, *Bioinformatics* (2001); *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids* (Durbin *et al.*, eds., 1999); *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins* (Baxeavanis & Oeullette eds., 1998)); Rashidi & Buehler, *Bioinformatics: Basic Applications in Biological Science and Medicine* (1999); *Introduction to Computational Molecular Biology* (Setubal *et al.*, eds 1997); *Bioinformatics: Methods and Protocols* (Misener & Krawetz, eds, 2000); *Bioinformatics: Sequence, Structure, and Databanks: A Practical Approach* (Higgins & Taylor, eds., 2000); Brown, *Bioinformatics: A Biologist's Guide to Biocomputing and the Internet* (2001); Han & Kamber, *Data Mining: Concepts and Techniques* (2000); and Waterman, *Introduction to Computational Biology: Maps, Sequences, and Genomes* (1995).

The present invention provides a computer database comprising a computer and software for storing in computer-retrievable form assay data records cross-tabulated, e.g., with data specifying the source of the target-containing sample from which each sequence specificity record was obtained.

In an exemplary embodiment, at least one of the sources of target-containing sample is from a control tissue sample known to be free of pathological disorders. In a variation, at least one of the sources is a known pathological tissue specimen, e.g., a neoplastic lesion or another tissue specimen to be analyzed for prostate cancer. In another variation, the assay records cross-tabulate one or more of the following parameters for each target species in a sample: (1) a unique identification code, which can include, e.g., a target molecular structure and/or characteristic separation coordinate (e.g., electrophoretic coordinates); (2) sample source; and (3) absolute and/or relative quantity of the target species present in the sample.

The invention also provides for the storage and retrieval of a collection of target data in a computer data storage apparatus, which can include magnetic disks, optical disks, magneto-optical disks, DRAM, SRAM, SGRAM, SDRAM, RDRAM, DDR RAM, magnetic bubble memory devices, and other data storage devices, including CPU registers and on-CPU data storage arrays. Typically, the target data records are stored as a bit pattern in an array of magnetic domains on a magnetizable medium or as an array of charge states or transistor gate states, such as an array of cells in a DRAM device (e.g., each cell comprised of a transistor and a charge storage area, which may be on the transistor). In one embodiment,

the invention provides such storage devices, and computer systems built therewith, comprising a bit pattern encoding a protein expression fingerprint record comprising unique identifiers for at least 10 target data records cross-tabulated with target source.

When the target is a peptide or nucleic acid, the invention preferably provides
5 a method for identifying related peptide or nucleic acid sequences, comprising performing a computerized comparison between a peptide or nucleic acid sequence assay record stored in or retrieved from a computer storage device or database and at least one other sequence. The comparison can include a sequence analysis or comparison algorithm or computer program embodiment thereof (e.g., FASTA, TFASTA, GAP, BESTFIT) and/or the comparison may
10 be of the relative amount of a peptide or nucleic acid sequence in a pool of sequences determined from a polypeptide or nucleic acid sample of a specimen.

The invention also preferably provides a magnetic disk, such as an IBM-compatible (DOS, Windows, Windows95/98/2000, Windows NT, OS/2) or other format (e.g., Linux, SunOS, Solaris, AIX, SCO Unix, VMS, MV, Macintosh, etc.) floppy diskette or
15 hard (fixed, Winchester) disk drive, comprising a bit pattern encoding data from an assay of the invention in a file format suitable for retrieval and processing in a computerized sequence analysis, comparison, or relative quantitation method.

The invention also provides a network, comprising a plurality of computing devices linked via a data link, such as an Ethernet cable (coax or 10BaseT), telephone line,
20 ISDN line, wireless network, optical fiber, or other suitable signal transmission medium, whereby at least one network device (e.g., computer, disk array, etc.) comprises a pattern of magnetic domains (e.g., magnetic disk) and/or charge domains (e.g., an array of DRAM cells) composing a bit pattern encoding data acquired from an assay of the invention.

The invention also provides a method for transmitting assay data that includes
25 generating an electronic signal on an electronic communications device, such as a modem, ISDN terminal adapter, DSL, cable modem, ATM switch, or the like, wherein the signal includes (in native or encrypted format) a bit pattern encoding data from an assay or a database comprising a plurality of assay results obtained by the method of the invention.

In a preferred embodiment, the invention provides a computer system for
30 comparing a query target to a database containing an array of data structures, such as an assay result obtained by the method of the invention, and ranking database targets based on the

degree of identity and gap weight to the target data. A central processor is preferably initialized to load and execute the computer program for alignment and/or comparison of the assay results. Data for a query target is entered into the central processor via an I/O device. Execution of the computer program results in the central processor retrieving the assay data from the data file, which comprises a binary description of an assay result.

The target data or record and the computer program can be transferred to secondary memory, which is typically random access memory (e.g., DRAM, SRAM, SGRAM, or SDRAM). Targets are ranked according to the degree of correspondence between a selected assay characteristic (e.g., binding to a selected affinity moiety) and the same characteristic of the query target and results are output via an I/O device. For example, a central processor can be a conventional computer (e.g., Intel Pentium, PowerPC, Alpha, PA-8000, SPARC, MIPS 4400, MIPS 10000, VAX, etc.); a program can be a commercial or public domain molecular biology software package (e.g., UWGCG Sequence Analysis Software, Darwin); a data file can be an optical or magnetic disk, a data server, a memory device (e.g., DRAM, SRAM, SGRAM, SDRAM, EPROM, bubble memory, flash memory, etc.); an I/O device can be a terminal comprising a video display and a keyboard, a modem, an ISDN terminal adapter, an Ethernet port, a punched card reader, a magnetic strip reader, or other suitable I/O device.

The invention also preferably provides the use of a computer system, such as that described above, which comprises: (1) a computer; (2) a stored bit pattern encoding a collection of peptide sequence specificity records obtained by the methods of the invention, which may be stored in the computer; (3) a comparison target, such as a query target; and (4) a program for alignment and comparison, typically with rank-ordering of comparison results on the basis of computed similarity values.

Characteristics of prostate cancer-associated proteins

Prostate cancer proteins of the present invention may be classified as secreted proteins, transmembrane proteins or intracellular proteins. In one embodiment, the prostate cancer protein is an intracellular protein. Intracellular proteins may be found in the cytoplasm and/or in the nucleus. Intracellular proteins are involved in all aspects of cellular function and replication (including, e.g., signaling pathways); aberrant expression of such

proteins often results in unregulated or disregulated cellular processes (*see, e.g., Molecular Biology of the Cell* (Alberts, ed., 3rd ed., 1994)). For example, many intracellular proteins have enzymatic activity such as protein kinase activity, protein phosphatase activity, protease activity, nucleotide cyclase activity, polymerase activity and the like. Intracellular proteins also serve as docking proteins that are involved in organizing complexes of proteins, or targeting proteins to various subcellular localizations, and are involved in maintaining the structural integrity of organelles.

An increasingly appreciated concept in characterizing proteins is the presence in the proteins of one or more motifs for which defined functions have been attributed. In addition to the highly conserved sequences found in the enzymatic domain of proteins, highly conserved sequences have been identified in proteins that are involved in protein-protein interaction. For example, Src-homology-2 (SH2) domains bind tyrosine-phosphorylated targets in a sequence dependent manner. PTB domains, which are distinct from SH2 domains, also bind tyrosine phosphorylated targets. SH3 domains bind to proline-rich targets. In addition, PH domains, tetratricopeptide repeats and WD domains to name only a few, have been shown to mediate protein-protein interactions. Some of these may also be involved in binding to phospholipids or other second messengers. As will be appreciated by one of ordinary skill in the art, these motifs can be identified on the basis of primary sequence; thus, an analysis of the sequence of proteins may provide insight into both the enzymatic potential of the molecule and/or molecules with which the protein may associate. One useful database is Pfam (protein families), which is a large collection of multiple sequence alignments and hidden Markov models covering many common protein domains. Versions are available via the internet from Washington University in St. Louis, the Sanger Center in England, and the Karolinska Institute in Sweden (*see, e.g., Bateman et al., Nuc. Acids Res.* 28:263-266 (2000); Sonnhammer *et al., Proteins* 28:405-420 (1997); Bateman *et al., Nuc. Acids Res.* 27:260-262 (1999); and Sonnhammer *et al., Nuc. Acids Res.* 26:320-322- (1998)).

In another embodiment, the prostate cancer sequences are transmembrane proteins. Transmembrane proteins are molecules that span a phospholipid bilayer of a cell. They may have an intracellular domain, an extracellular domain, or both. The intracellular domains of such proteins may have a number of functions including those already described

for intracellular proteins. For example, the intracellular domain may have enzymatic activity and/or may serve as a binding site for additional proteins. Frequently the intracellular domain of transmembrane proteins serves both roles. For example certain receptor tyrosine kinases have both protein kinase activity and SH2 domains. In addition, autophosphorylation of tyrosines on the receptor molecule itself, creates binding sites for additional SH2 domain containing proteins.

Transmembrane proteins may contain from one to many transmembrane domains. For example, receptor tyrosine kinases, certain cytokine receptors, receptor guanylyl cyclases and receptor serine/threonine protein kinases contain a single transmembrane domain. However, various other proteins including channels and adenylyl cyclases contain numerous transmembrane domains. Many important cell surface receptors such as G protein coupled receptors (GPCRs) are classified as “seven transmembrane domain” proteins, as they contain 7 membrane spanning regions. Characteristics of transmembrane domains include approximately 20 consecutive hydrophobic amino acids that may be followed by charged amino acids. Therefore, upon analysis of the amino acid sequence of a particular protein, the localization and number of transmembrane domains within the protein may be predicted (*see, e.g.* PSORT web site <http://psort.nibb.ac.jp/>). Important transmembrane protein receptors include, but are not limited to the insulin receptor, insulin-like growth factor receptor, human growth hormone receptor, glucose transporters, transferrin receptor, epidermal growth factor receptor, low density lipoprotein receptor, epidermal growth factor receptor, leptin receptor, interleukin receptors, e.g. IL-1 receptor, IL-2 receptor,

The extracellular domains of transmembrane proteins are diverse; however, conserved motifs are found repeatedly among various extracellular domains. Conserved structure and/or functions have been ascribed to different extracellular motifs. Many extracellular domains are involved in binding to other molecules. In one aspect, extracellular domains are found on receptors. Factors that bind the receptor domain include circulating ligands, which may be peptides, proteins, or small molecules such as adenosine and the like. For example, growth factors such as EGF, FGF and PDGF are circulating growth factors that bind to their cognate receptors to initiate a variety of cellular responses. Other factors include cytokines, mitogenic factors, neurotrophic factors and the like. Extracellular domains also

bind to cell-associated molecules. In this respect, they mediate cell-cell interactions. Cell-associated ligands can be tethered to the cell, e.g., via a glycosylphosphatidylinositol (GPI) anchor, or may themselves be transmembrane proteins. Extracellular domains also associate with the extracellular matrix and contribute to the maintenance of the cell structure.

5 Prostate cancer proteins that are transmembrane are particularly preferred in the present invention as they are readily accessible targets for immunotherapeutics, as are described herein. In addition, as outlined below, transmembrane proteins can be also useful in imaging modalities. Antibodies may be used to label such readily accessible proteins *in situ*. Alternatively, antibodies can also label intracellular proteins, in which case samples are
10 typically permeablized to provide access to intracellular proteins.

It will also be appreciated by those in the art that a transmembrane protein can be made soluble by removing transmembrane sequences, e.g., through recombinant methods. Furthermore, transmembrane proteins that have been made soluble can be made to be secreted through recombinant means by adding an appropriate signal sequence.

15 In another embodiment, the prostate cancer proteins are secreted proteins; the secretion of which can be either constitutive or regulated. These proteins have a signal peptide or signal sequence that targets the molecule to the secretory pathway. Secreted proteins are involved in numerous physiological events; by virtue of their circulating nature, they serve to transmit signals to various other cell types. The secreted protein may function in
20 an autocrine manner (acting on the cell that secreted the factor), a paracrine manner (acting on cells in close proximity to the cell that secreted the factor) or an endocrine manner (acting on cells at a distance). Thus secreted molecules find use in modulating or altering numerous aspects of physiology. Prostate cancer proteins that are secreted proteins are particularly preferred in the present invention as they serve as good targets for diagnostic markers, e.g.,
25 for blood, plasma, serum, or stool tests.

Use of prostate cancer nucleic acids

As described above, prostate cancer sequence is initially identified by substantial nucleic acid and/or amino acid sequence homology or linkage to the prostate
30 cancer sequences outlined herein. Such homology can be based upon the overall nucleic acid or amino acid sequence, and is generally determined as outlined below, using either

homology programs or hybridization conditions. Typically, linked sequences on a mRNA are found on the same molecule.

The prostate cancer nucleic acid sequences of the invention, e.g., the sequences in Tables 1-16, can be fragments of larger genes, i.e., they are nucleic acid segments. “Genes” in this context includes coding regions, non-coding regions, and mixtures of coding and non-coding regions. Accordingly, as will be appreciated by those in the art, using the sequences provided herein, extended sequences, in either direction, of the prostate cancer genes can be obtained, using techniques well known in the art for cloning either longer sequences or the full length sequences; see Ausubel, *et al.*, *supra*. Much can be done by informatics and many sequences can be clustered to include multiple sequences corresponding to a single gene, e.g., systems such as UniGene (see, <http://www.ncbi.nlm.nih.gov/UniGene/>).

Once the prostate cancer nucleic acid is identified, it can be cloned and, if necessary, its constituent parts recombined to form the entire prostate cancer nucleic acid coding regions or the entire mRNA sequence. Once isolated from its natural source, e.g., contained within a plasmid or other vector or excised therefrom as a linear nucleic acid segment, the recombinant prostate cancer nucleic acid can be further-used as a probe to identify and isolate other prostate cancer nucleic acids, e.g., extended coding regions. It can also be used as a “precursor” nucleic acid to make modified or variant prostate cancer nucleic acids and proteins.

The prostate cancer nucleic acids of the present invention are used in several ways. In a first embodiment, nucleic acid probes to the prostate cancer nucleic acids are made and attached to biochips to be used in screening and diagnostic methods, as outlined below, or for administration, e.g., for gene therapy, vaccine, and/or antisense applications. Alternatively, the prostate cancer nucleic acids that include coding regions of prostate cancer proteins can be put into expression vectors for the expression of prostate cancer proteins, again for screening purposes or for administration to a patient.

In a preferred embodiment, nucleic acid probes to prostate cancer nucleic acids (both the nucleic acid sequences outlined in the figures and/or the complements thereof) are made. The nucleic acid probes attached to the biochip are designed to be substantially complementary to the prostate cancer nucleic acids, *i.e.* the target sequence (either the target

sequence of the sample or to other probe sequences, e.g., in sandwich assays), such that hybridization of the target sequence and the probes of the present invention occurs. As outlined below, this complementarity need not be perfect; there may be any number of base pair mismatches which will interfere with hybridization between the target sequence and the single stranded nucleic acids of the present invention. However, if the number of mutations is so great that no hybridization can occur under even the least stringent of hybridization conditions, the sequence is not a complementary target sequence. Thus, by “substantially complementary” herein is meant that the probes are sufficiently complementary to the target sequences to hybridize under normal reaction conditions, particularly high stringency conditions, as outlined herein.

A nucleic acid probe is generally single stranded but can be partially single and partially double stranded. The strandedness of the probe is dictated by the structure, composition, and properties of the target sequence. In general, the nucleic acid probes range from about 8 to about 100 bases long, with from about 10 to about 80 bases being preferred, and from about 30 to about 50 bases being particularly preferred. That is, generally whole genes are not used. In some embodiments, much longer nucleic acids can be used, up to hundreds of bases.

In a preferred embodiment, more than one probe per sequence is used, with either overlapping probes or probes to different sections of the target being used. That is, two, three, four or more probes, with three being preferred, are used to build in a redundancy for a particular target. The probes can be overlapping (i.e., have some sequence in common), or separate. In some cases, PCR primers may be used to amplify signal for higher sensitivity.

As will be appreciated by those in the art, nucleic acids can be attached or immobilized to a solid support in a wide variety of ways. By “immobilized” and grammatical equivalents herein is meant the association or binding between the nucleic acid probe and the solid support is sufficient to be stable under the conditions of binding, washing, analysis, and removal as outlined below. The binding can typically be covalent or non-covalent. By “non-covalent binding” and grammatical equivalents herein is meant one or more of electrostatic, hydrophilic, and hydrophobic interactions. Included in non-covalent binding is the covalent attachment of a molecule, such as, streptavidin to the support and the non-covalent binding of the biotinylated probe to the streptavidin. By “covalent binding” and grammatical

equivalents herein is meant that the two moieties, the solid support and the probe, are attached by at least one bond, including sigma bonds, pi bonds and coordination bonds. Covalent bonds can be formed directly between the probe and the solid support or can be formed by a cross linker or by inclusion of a specific reactive group on either the solid support or the probe or both molecules. Immobilization may also involve a combination of covalent and non-covalent interactions.

In general, the probes are attached to the biochip in a wide variety of ways, as will be appreciated by those in the art. As described herein, the nucleic acids can either be synthesized first, with subsequent attachment to the biochip, or can be directly synthesized on the biochip.

The biochip comprises a suitable solid substrate. By "substrate" or "solid support" or other grammatical equivalents herein is meant a material that can be modified to contain discrete individual sites appropriate for the attachment or association of the nucleic acid probes and is amenable to at least one detection method. As will be appreciated by those in the art, the number of possible substrates are very large, and include, but are not limited to, glass and modified or functionalized glass, plastics (including acrylics, polystyrene and copolymers of styrene and other materials, polypropylene, polyethylene, polybutylene, polyurethanes, Teflon, etc.), polysaccharides, nylon or nitrocellulose, resins, silica or silica-based materials including silicon and modified silicon, carbon, metals, inorganic glasses, plastics, etc. In general, the substrates allow optical detection and do not appreciably fluoresce. A preferred substrate is described in copending application entitled Reusable Low Fluorescent Plastic Biochip, U.S. Application Serial No. 09/270,214, filed March 15, 1999, herein incorporated by reference in its entirety.

Generally the substrate is planar, although as will be appreciated by those in the art, other configurations of substrates may be used as well. For example, the probes may be placed on the inside surface of a tube, for flow-through sample analysis to minimize sample volume. Similarly, the substrate may be flexible, such as a flexible foam, including closed cell foams made of particular plastics.

In a preferred embodiment, the surface of the biochip and the probe may be derivatized with chemical functional groups for subsequent attachment of the two. Thus, e.g., the biochip is derivatized with a chemical functional group including, but not limited to,

amino groups, carboxy groups, oxo groups and thiol groups, with amino groups being particularly preferred. Using these functional groups, the probes can be attached using functional groups on the probes. For example, nucleic acids containing amino groups can be attached to surfaces comprising amino groups, e.g. using linkers as are known in the art; e.g.,
5 homo-or hetero-bifunctional linkers as are well known (*see* 1994 Pierce Chemical Company catalog, technical section on cross-linkers, pages 155-200). In addition, in some cases, additional linkers, such as alkyl groups (including substituted and heteroalkyl groups) may be used.

In this embodiment, oligonucleotides are synthesized as is known in the art,
10 and then attached to the surface of the solid support. As will be appreciated by those skilled in the art, either the 5' or 3' terminus may be attached to the solid support, or attachment may be via an internal nucleoside.

In another embodiment, the immobilization to the solid support may be very strong, yet non-covalent. For example, biotinylated oligonucleotides can be made, which
15 bind to surfaces covalently coated with streptavidin, resulting in attachment.

Alternatively, the oligonucleotides may be synthesized on the surface, as is known in the art. For example, photoactivation techniques utilizing photopolymerization compounds and techniques are used. In a preferred embodiment, the nucleic acids can be synthesized in situ, using well known photolithographic techniques, such as those described
20 in WO 95/25116; WO 95/35505; U.S. Patent Nos. 5,700,637 and 5,445,934; and references cited within, all of which are expressly incorporated by reference; these methods of attachment form the basis of the Affimetrix GeneChip™ technology.

Often, amplification-based assays are performed to measure the expression level of prostate cancer-associated sequences. These assays are typically performed in
25 conjunction with reverse transcription. In such assays, a prostate cancer-associated nucleic acid sequence acts as a template in an amplification reaction (e.g., Polymerase Chain Reaction, or PCR). In a quantitative amplification, the amount of amplification product will be proportional to the amount of template in the original sample. Comparison to appropriate controls provides a measure of the amount of prostate cancer-associated RNA. Methods of
30 quantitative amplification are well known to those of skill in the art. Detailed protocols for

quantitative PCR are provided, e.g., in Innis *et al.*, *PCR Protocols, A Guide to Methods and Applications* (1990).

In some embodiments, a TaqMan based assay is used to measure expression. TaqMan based assays use a fluorogenic oligonucleotide probe that contains a 5' fluorescent dye and a 3' quenching agent. The probe hybridizes to a PCR product, but cannot itself be extended due to a blocking agent at the 3' end. When the PCR product is amplified in subsequent cycles, the 5' nuclease activity of the polymerase, e.g., AmpliTaq, results in the cleavage of the TaqMan probe. This cleavage separates the 5' fluorescent dye and the 3' quenching agent, thereby resulting in an increase in fluorescence as a function of

amplification (*see*, e.g., literature provided by Perkin-Elmer, e.g., www2.perkin-elmer.com).

Other suitable amplification methods include, but are not limited to, ligase chain reaction (LCR) (*see* Wu & Wallace, *Genomics* 4:560 (1989), Landegren *et al.*, *Science* 241:1077 (1988), and Barringer *et al.*, *Gene* 89:117 (1990)), transcription amplification (Kwoh *et al.*, *Proc. Natl. Acad. Sci. USA* 86:1173 (1989)), self-sustained sequence replication (Guatelli *et al.*, *Proc. Nat. Acad. Sci. USA* 87:1874 (1990)), dot PCR, and linker adapter PCR, etc.

Expression of prostate cancer proteins from nucleic acids

In a preferred embodiment, prostate cancer nucleic acids, e.g., encoding prostate cancer proteins are used to make a variety of expression vectors to express prostate cancer proteins which can then be used in screening assays, as described below. Expression vectors and recombinant DNA technology are well known to those of skill in the art (*see*, e.g., Ausubel, *supra*, and *Gene Expression Systems* (Fernandez & Hoeffler, eds, 1999)) and are used to express proteins. The expression vectors may be either self-replicating extrachromosomal vectors or vectors which integrate into a host genome. Generally, these expression vectors include transcriptional and translational regulatory nucleic acid operably linked to the nucleic acid encoding the prostate cancer protein. The term "control sequences" refers to DNA sequences used for the expression of an operably linked coding sequence in a particular host organism. Control sequences that are suitable for prokaryotes, e.g., include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

Nucleic acid is “operably linked” when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, “operably linked” means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is typically accomplished by ligation at convenient restriction sites. If such sites do not exist, synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice. Transcriptional and translational regulatory nucleic acid will generally be appropriate to the host cell used to express the prostate cancer protein. Numerous types of appropriate expression vectors, and suitable regulatory sequences are known in the art for a variety of host cells.

In general, transcriptional and translational regulatory sequences may include, but are not limited to, promoter sequences, ribosomal binding sites, transcriptional start and stop sequences, translational start and stop sequences, and enhancer or activator sequences. In a preferred embodiment, the regulatory sequences include a promoter and transcriptional start and stop sequences.

Promoter sequences encode either constitutive or inducible promoters. The promoters may be either naturally occurring promoters or hybrid promoters. Hybrid promoters, which combine elements of more than one promoter, are also known in the art, and are useful in the present invention.

In addition, an expression vector may comprise additional elements. For example, the expression vector may have two replication systems, thus allowing it to be maintained in two organisms, e.g. in mammalian or insect cells for expression and in a procaryotic host for cloning and amplification. Furthermore, for integrating expression vectors, the expression vector contains at least one sequence homologous to the host cell genome, and preferably two homologous sequences which flank the expression construct.

The integrating vector may be directed to a specific locus in the host cell by selecting the

appropriate homologous sequence for inclusion in the vector. Constructs for integrating vectors are well known in the art (e.g., Fernandez & Hoeffler, *supra*).

In addition, in a preferred embodiment, the expression vector contains a selectable marker gene to allow the selection of transformed host cells. Selection genes are well known in the art and will vary with the host cell used.

The prostate cancer proteins of the present invention are produced by culturing a host cell transformed with an expression vector containing nucleic acid encoding a prostate cancer protein, under the appropriate conditions to induce or cause expression of the prostate cancer protein. Conditions appropriate for prostate cancer protein expression will vary with the choice of the expression vector and the host cell, and will be easily ascertained by one skilled in the art through routine experimentation or optimization. For example, the use of constitutive promoters in the expression vector will require optimizing the growth and proliferation of the host cell, while the use of an inducible promoter requires the appropriate growth conditions for induction. In addition, in some embodiments, the timing of the harvest is important. For example, the baculoviral systems used in insect cell expression are lytic viruses, and thus harvest time selection can be crucial for product yield.

Appropriate host cells include yeast, bacteria, archaeobacteria, fungi, and insect and animal cells, including mammalian cells. Of particular interest are *Saccharomyces cerevisiae* and other yeasts, *E. coli*, *Bacillus subtilis*, Sf9 cells, C129 cells, 293 cells, *Neurospora*, BHK, CHO, COS, HeLa cells, HUVEC (human umbilical vein endothelial cells), THP1 cells (a macrophage cell line) and various other human cells and cell lines.

In a preferred embodiment, the prostate cancer proteins are expressed in mammalian cells. Mammalian expression systems are also known in the art, and include retroviral and adenoviral systems. One expression vector system is a retroviral vector system such as is generally described in PCT/US97/01019 and PCT/US97/01048, both of which are hereby expressly incorporated by reference. Of particular use as mammalian promoters are the promoters from mammalian viral genes, since the viral genes are often highly expressed and have a broad host range. Examples include the SV40 early promoter, mouse mammary tumor virus LTR promoter, adenovirus major late promoter, herpes simplex virus promoter, and the CMV promoter (*see, e.g.,* Fernandez & Hoeffler, *supra*). Typically, transcription termination and polyadenylation sequences recognized by mammalian cells are regulatory

regions located 3' to the translation stop codon and thus, together with the promoter elements, flank the coding sequence. Examples of transcription terminator and polyadenylation signals include those derived from SV40.

The methods of introducing exogenous nucleic acid into mammalian hosts, as well as other hosts, is well known in the art, and will vary with the host cell used. Techniques include dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, viral infection, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei.

In a preferred embodiment, prostate cancer proteins are expressed in bacterial systems. Bacterial expression systems are well known in the art. Promoters from bacteriophage may also be used and are known in the art. In addition, synthetic promoters and hybrid promoters are also useful; e.g., the tac promoter is a hybrid of the trp and lac promoter sequences. Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. In addition to a functioning promoter sequence, an efficient ribosome binding site is desirable. The expression vector may also include a signal peptide sequence that provides for secretion of the prostate cancer protein in bacteria. The protein is either secreted into the growth media (gram-positive bacteria) or into the periplasmic space, located between the inner and outer membrane of the cell (gram-negative bacteria). The bacterial expression vector may also include a selectable marker gene to allow for the selection of bacterial strains that have been transformed. Suitable selection genes include genes which render the bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin, neomycin and tetracycline. Selectable markers also include biosynthetic genes, such as those in the histidine, tryptophan and leucine biosynthetic pathways. These components are assembled into expression vectors. Expression vectors for bacteria are well known in the art, and include vectors for *Bacillus subtilis*, *E. coli*, *Streptococcus cremoris*, and *Streptococcus lividans*, among others (e.g., Fernandez & Hoeffler, *supra*). The bacterial expression vectors are transformed into bacterial host cells using techniques well known in the art, such as calcium chloride treatment, electroporation, and others.

In one embodiment, prostate cancer proteins are produced in insect cells. Expression vectors for the transformation of insect cells, and in particular, baculovirus-based expression vectors, are well known in the art.

In a preferred embodiment, prostate cancer protein is produced in yeast cells.

5 Yeast expression systems are well known in the art, and include expression vectors for *Saccharomyces cerevisiae*, *Candida albicans* and *C. maltosa*, *Hansenula polymorpha*, *Kluyveromyces fragilis* and *K. lactis*, *Pichia guillerimondii* and *P. pastoris*, *Schizosaccharomyces pombe*, and *Yarrowia lipolytica*.

10 The prostate cancer protein may also be made as a fusion protein, using techniques well known in the art. Thus, e.g., for the creation of monoclonal antibodies, if the desired epitope is small, the prostate cancer protein may be fused to a carrier protein to form an immunogen. Alternatively, the prostate cancer protein may be made as a fusion protein to increase expression, or for other reasons. For example, when the prostate cancer protein is a prostate cancer peptide, the nucleic acid encoding the peptide may be linked to other nucleic
15 acid for expression purposes.

In a preferred embodiment, the prostate cancer protein is purified or isolated after expression. Prostate cancer proteins may be isolated or purified in a variety of ways known to those skilled in the art depending on what other components are present in the sample. Standard purification methods include electrophoretic, molecular, immunological
20 and chromatographic techniques, including ion exchange, hydrophobic, affinity, and reverse-phase HPLC chromatography, and chromatofocusing. For example, the prostate cancer protein may be purified using a standard anti-prostate cancer protein antibody column. Ultrafiltration and diafiltration techniques, in conjunction with protein concentration, are also useful. For general guidance in suitable purification techniques, see Scopes, *Protein*
25 *Purification* (1982). The degree of purification necessary will vary depending on the use of the prostate cancer protein. In some instances no purification will be necessary.

Once expressed and purified if necessary, the prostate cancer proteins and nucleic acids are useful in a number of applications. They may be used as immunoselection reagents, as vaccine reagents, as screening agents, etc.

Variants of prostate cancer proteins

In one embodiment, the prostate cancer proteins are derivative or variant prostate cancer proteins as compared to the wild-type sequence. That is, as outlined more fully below, the derivative prostate cancer peptide will often contain at least one amino acid substitution, deletion or insertion, with amino acid substitutions being particularly preferred. The amino acid substitution, insertion or deletion may occur at any residue within the prostate cancer peptide.

Also included within one embodiment of prostate cancer proteins of the present invention are amino acid sequence variants. These variants typically fall into one or more of three classes: substitutional, insertional or deletional variants. These variants ordinarily are prepared by site specific mutagenesis of nucleotides in the DNA encoding the prostate cancer protein, using cassette or PCR mutagenesis or other techniques well known in the art, to produce DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture as outlined above. However, variant prostate cancer protein fragments having up to about 100-150 residues may be prepared by in vitro synthesis using established techniques. Amino acid sequence variants are characterized by the predetermined nature of the variation, a feature that sets them apart from naturally occurring allelic or interspecies variation of the prostate cancer protein amino acid sequence. The variants typically exhibit the same qualitative biological activity as the naturally occurring analogue, although variants can also be selected which have modified characteristics as will be more fully outlined below.

While the site or region for introducing an amino acid sequence variation is predetermined, the mutation per se need not be predetermined. For example, in order to optimize the performance of a mutation at a given site, random mutagenesis may be conducted at the target codon or region and the expressed prostate cancer variants screened for the optimal combination of desired activity. Techniques for making substitution mutations at predetermined sites in DNA having a known sequence are well known, e.g., M13 primer mutagenesis and PCR mutagenesis. Screening of the mutants is done using assays of prostate cancer protein activities.

Amino acid substitutions are typically of single residues; insertions usually will be on the order of from about 1 to 20 amino acids, although considerably larger

insertions may be tolerated. Deletions range from about 1 to about 20 residues, although in some cases deletions may be much larger.

Substitutions, deletions, insertions or any combination thereof may be used to arrive at a final derivative. Generally these changes are done on a few amino acids to minimize the alteration of the molecule. However, larger changes may be tolerated in certain circumstances. When small alterations in the characteristics of the prostate cancer protein are desired, substitutions are generally made in accordance with the amino acid substitution relationships provided in the definition section.

The variants typically exhibit the same qualitative biological activity and will elicit the same immune response as the naturally-occurring analog, although variants also are selected to modify the characteristics of the prostate cancer proteins as needed. Alternatively, the variant may be designed such that the biological activity of the prostate cancer protein is altered. For example, glycosylation sites may be altered or removed.

Substantial changes in function or immunological identity are made by selecting substitutions that are less conservative than those described above. For example, substitutions may be made which more significantly affect: the structure of the polypeptide backbone in the area of the alteration, for example the alpha-helical or beta-sheet structure; the charge or hydrophobicity of the molecule at the target site; or the bulk of the side chain. The substitutions which in general are expected to produce the greatest changes in the polypeptide's properties are those in which (a) a hydrophilic residue, e.g. seryl or threonyl is substituted for (or by) a hydrophobic residue, e.g. leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue having an electropositive side chain, e.g. lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g. glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g. phenylalanine, is substituted for (or by) one not having a side chain, e.g. glycine.

Covalent modifications of prostate cancer polypeptides are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a prostate cancer polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N-or C-terminal residues of a prostate cancer polypeptide. Derivatization with bifunctional agents is useful, for instance, for crosslinking prostate cancer polypeptides to a water-insoluble support matrix or surface for

use in the method for purifying anti-prostate cancer polypeptide antibodies or screening assays, as is more fully described below. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, e.g., esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl
5 esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-((p-azidophenyl)dithio)propioimide.

Other modifications include deamidation of glutamyl and asparaginy residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl, threonyl or tyrosyl residues,
10 methylation of the amino groups of the lysine, arginine, and histidine side chains (Creighton, *Proteins: Structure and Molecular Properties*, pp. 79-86 (1983)), acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the prostate cancer polypeptide included within the scope of this invention comprises altering the native glycosylation pattern
15 of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence prostate cancer polypeptide, and/or adding one or more glycosylation sites that are not present in the native sequence prostate cancer polypeptide. Glycosylation patterns can be altered in many ways. For example the use of different cell types to express prostate cancer-associated
20 sequences can result in different glycosylation patterns.

Addition of glycosylation sites to prostate cancer polypeptides may also be accomplished by altering the amino acid sequence thereof. The alteration may be made, e.g., by the addition of, or substitution by, one or more serine or threonine residues to the native sequence prostate cancer polypeptide (for O-linked glycosylation sites). The prostate cancer
25 amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the prostate cancer polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the prostate cancer polypeptide is by chemical or enzymatic coupling of glycosides to the
30 polypeptide. Such methods are described in the art, e.g., in WO 87/05330, and in Aplin & Wriston, *CRC Crit. Rev. Biochem.*, pp. 259-306 (1981).

Removal of carbohydrate moieties present on the prostate cancer polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin,
5 *et al.*, *Arch. Biochem. Biophys.*, 259:52 (1987) and by Edge *et al.*, *Anal. Biochem.*, 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo-and exo-glycosidases as described by Thotakura *et al.*, *Meth. Enzymol.*, 138:350 (1987).

Another type of covalent modification of prostate cancer comprises linking the
10 prostate cancer polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol, polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

Prostate cancer polypeptides of the present invention may also be modified in a way to form chimeric molecules comprising a prostate cancer polypeptide fused to another,
15 heterologous polypeptide or amino acid sequence. In one embodiment, such a chimeric molecule comprises a fusion of a prostate cancer polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino-or carboxyl-terminus of the prostate cancer polypeptide. The presence of such epitope-tagged forms of a prostate cancer polypeptide can be detected using
20 an antibody against the tag polypeptide. Also, provision of the epitope tag enables the prostate cancer polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. In an alternative embodiment, the chimeric molecule may comprise a fusion of a prostate cancer polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of
25 the chimeric molecule, such a fusion could be to the Fc region of an IgG molecule.

Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; HIS6 and metal chelation tags, the flu HA tag polypeptide and its antibody 12CA5 (Field *et al.*, *Mol. Cell. Biol.* 8:2159-2165 (1988)); the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and
30 9E10 antibodies thereto (Evan *et al.*, *Molecular and Cellular Biology* 5:3610-3616 (1985)); and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody (Paborsky *et al.*,

Protein Engineering 3(6):547-553 (1990)). Other tag polypeptides include the Flag-peptide (Hopp *et al.*, *BioTechnology* 6:1204-1210 (1988)); the KT3 epitope peptide (Martin *et al.*, *Science* 255:192-194 (1992)); tubulin epitope peptide (Skinner *et al.*, *J. Biol. Chem.* 266:15163-15166 (1991)); and the T7 gene 10 protein peptide tag (Lutz-Freyermuth *et al.*,
5 *Proc. Natl. Acad. Sci. USA* 87:6393-6397 (1990)).

Also included are other prostate cancer proteins of the prostate cancer family, and prostate cancer proteins from other organisms, which are cloned and expressed as outlined below. Thus, probe or degenerate polymerase chain reaction (PCR) primer sequences may be used to find other related prostate cancer proteins from humans or other
10 organisms. As will be appreciated by those in the art, particularly useful probe and/or PCR primer sequences include the unique areas of the prostate cancer nucleic acid sequence. As is generally known in the art, preferred PCR primers are from about 15 to about 35 nucleotides in length, with from about 20 to about 30 being preferred, and may contain inosine as needed. The conditions for the PCR reaction are well known in the art (e.g., Innis, PCR Protocols,
15 *supra*).

Antibodies to prostate cancer proteins

In a preferred embodiment, when the prostate cancer protein is to be used to generate antibodies, e.g., for immunotherapy or immunodiagnosis, the prostate cancer protein
20 should share at least one epitope or determinant with the full length protein. By "epitope" or "determinant" herein is typically meant a portion of a protein which will generate and/or bind an antibody or T-cell receptor in the context of MHC. Thus, in most instances, antibodies made to a smaller prostate cancer protein will be able to bind to the full-length protein, particularly linear epitopes. In a preferred embodiment; the epitope is unique; that is,
25 antibodies generated to a unique epitope show little or no cross-reactivity.

Methods of preparing polyclonal antibodies are known to the skilled artisan (e.g., Coligan, *supra*; and Harlow & Lane, *supra*). Polyclonal antibodies can be raised in a mammal, e.g., by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple
30 subcutaneous or intraperitoneal injections. The immunizing agent may include a protein encoded by a nucleic acid of the figures or fragment thereof or a fusion protein thereof. It

may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete
5 adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

The antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler &
10 Milstein, *Nature* 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized in vitro. The immunizing agent will typically include a polypeptide encoded by a nucleic acid of Tables 1-
15 16 fragment thereof, or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, *Monoclonal Antibodies: Principles and Practice*,
20 pp. 59-103 (1986)). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme
25 hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

In one embodiment, the antibodies are bispecific antibodies. Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding
30 specificities for at least two different antigens or that have binding specificities for two epitopes on the same antigen. In one embodiment, one of the binding specificities is for a

protein encoded by a nucleic acid Tables 1-16 or a fragment thereof, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit, preferably one that is tumor specific. Alternatively, tetramer-type technology may create multivalent reagents.

5 In a preferred embodiment, the antibodies to prostate cancer protein are capable of reducing or eliminating a biological function of a prostate cancer protein, as is described below. That is, the addition of anti-prostate cancer protein antibodies (either polyclonal or preferably monoclonal) to prostate cancer tissue (or cells containing prostate cancer) may reduce or eliminate the prostate cancer. Generally, at least a 25% decrease in
10 activity, growth, size or the like is preferred, with at least about 50% being particularly preferred and about a 95-100% decrease being especially preferred.

 In a preferred embodiment the antibodies to the prostate cancer proteins are humanized antibodies (e.g., Xenex Biosciences, Mederex, Inc., Abgenix, Inc., Protein Design Labs, Inc.) Humanized forms of non-human (e.g., murine) antibodies are chimeric
15 molecules of immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-
20 human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, a humanized antibody will comprise
25 substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework (FR) regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human
30 immunoglobulin (Jones *et al.*, *Nature* 321:522-525 (1986); Riechmann *et al.*, *Nature* 332:323-329 (1988); and Presta, *Curr. Op. Struct. Biol.* 2:593-596 (1992)). Humanization

can be essentially performed following the method of Winter and co-workers (Jones *et al.*, *Nature* 321:522-525 (1986); Riechmann *et al.*, *Nature* 332:323-327 (1988); Verhoeyen *et al.*, *Science* 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such humanized antibodies are
5 chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species.

Human antibodies can also be produced using various techniques known in the art, including phage display libraries (Hoogenboom & Winter, *J. Mol. Biol.* 227:381 (1991);
10 Marks *et al.*, *J. Mol. Biol.* 222:581 (1991)). The techniques of Cole *et al.* and Boerner *et al.* are also available for the preparation of human monoclonal antibodies (Cole *et al.*, *Monoclonal Antibodies and Cancer Therapy*, p. 77 (1985) and Boerner *et al.*, *J. Immunol.* 147(1):86-95 (1991)). Similarly, human antibodies can be made by introducing of human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous
15 immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, e.g., in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in the following scientific publications: Marks *et al.*, *Bio/Technology* 10:779-
20 783 (1992); Lonberg *et al.*, *Nature* 368:856-859 (1994); Morrison, *Nature* 368:812-13 (1994); Fishwild *et al.*, *Nature Biotechnology* 14:845-51 (1996); Neuberger, *Nature Biotechnology* 14:826 (1996); Lonberg & Huszar, *Intern. Rev. Immunol.* 13:65-93 (1995).

By immunotherapy is meant treatment of prostate cancer with an antibody raised against prostate cancer proteins. As used herein, immunotherapy can be passive or
25 active. Passive immunotherapy as defined herein is the passive transfer of antibody to a recipient (patient). Active immunization is the induction of antibody and/or T-cell responses in a recipient (patient). Induction of an immune response is the result of providing the recipient with an antigen to which antibodies are raised. As appreciated by one of ordinary skill in the art, the antigen may be provided by injecting a polypeptide against which
30 antibodies are desired to be raised into a recipient, or contacting the recipient with a nucleic

acid capable of expressing the antigen and under conditions for expression of the antigen, leading to an immune response.

In a preferred embodiment the prostate cancer proteins against which antibodies are raised are secreted proteins as described above. Without being bound by theory, antibodies used for treatment, bind and prevent the secreted protein from binding to its receptor, thereby inactivating the secreted prostate cancer protein.

In another preferred embodiment, the prostate cancer protein to which antibodies are raised is a transmembrane protein. Without being bound by theory, antibodies used for treatment, bind the extracellular domain of the prostate cancer protein and prevent it from binding to other proteins, such as circulating ligands or cell-associated molecules. The antibody may cause down-regulation of the transmembrane prostate cancer protein. As will be appreciated by one of ordinary skill in the art, the antibody may be a competitive, non-competitive or uncompetitive inhibitor of protein binding to the extracellular domain of the prostate cancer protein. The antibody is also an antagonist of the prostate cancer protein.

Further, the antibody prevents activation of the transmembrane prostate cancer protein. In one aspect, when the antibody prevents the binding of other molecules to the prostate cancer protein, the antibody prevents growth of the cell. The antibody may also be used to target or sensitize the cell to cytotoxic agents, including, but not limited to TNF- α , TNF- β , IL-1, INF- γ and IL-2, or chemotherapeutic agents including 5FU, vinblastine, actinomycin D, cisplatin, methotrexate, and the like. In some instances the antibody belongs to a sub-type that activates serum complement when complexed with the transmembrane protein thereby mediating cytotoxicity or antigen-dependent cytotoxicity (ADCC). Thus, prostate cancer is treated by administering to a patient antibodies directed against the transmembrane prostate cancer protein. Antibody-labeling may activate a co-toxin, localize a toxin payload, or otherwise provide means to locally ablate cells.

In another preferred embodiment, the antibody is conjugated to an effector moiety. The effector moiety can be any number of molecules, including labelling moieties such as radioactive labels or fluorescent labels, or can be a therapeutic moiety. In one aspect the therapeutic moiety is a small molecule that modulates the activity of the prostate cancer protein. In another aspect the therapeutic moiety modulates the activity of molecules associated with or in close proximity to the prostate cancer protein. The therapeutic moiety

may inhibit enzymatic activity such as protease or collagenase or protein kinase activity associated with prostate cancer.

In a preferred embodiment, the therapeutic moiety can also be a cytotoxic agent. In this method, targeting the cytotoxic agent to prostate cancer tissue or cells, results in a reduction in the number of afflicted cells, thereby reducing symptoms associated with prostate cancer. Cytotoxic agents are numerous and varied and include, but are not limited to, cytotoxic drugs or toxins or active fragments of such toxins. Suitable toxins and their corresponding fragments include diphtheria A chain, exotoxin A chain, ricin A chain, abrin A chain, curcin, crotin, phenomycin, enomycin and the like. Cytotoxic agents also include radiochemicals made by conjugating radioisotopes to antibodies raised against prostate cancer proteins, or binding of a radionuclide to a chelating agent that has been covalently attached to the antibody. Targeting the therapeutic moiety to transmembrane prostate cancer proteins not only serves to increase the local concentration of therapeutic moiety in the prostate cancer afflicted area, but also serves to reduce deleterious side effects that may be associated with the therapeutic moiety.

In another preferred embodiment, the prostate cancer protein against which the antibodies are raised is an intracellular protein. In this case, the antibody may be conjugated to a protein which facilitates entry into the cell. In one case, the antibody enters the cell by endocytosis. In another embodiment, a nucleic acid encoding the antibody is administered to the individual or cell. Moreover, wherein the prostate cancer protein can be targeted within a cell, i.e., the nucleus, an antibody thereto contains a signal for that target localization, i.e., a nuclear localization signal.

The prostate cancer antibodies of the invention specifically bind to prostate cancer proteins. By "specifically bind" herein is meant that the antibodies bind to the protein with a K_d of at least about 0.1 mM, more usually at least about 1 μ M, preferably at least about 0.1 μ M or better, and most preferably, 0.01 μ M or better. Selectivity of binding is also important.

Detection of prostate cancer sequence for diagnostic and therapeutic applications

In one aspect, the RNA expression levels of genes are determined for different cellular states in the prostate cancer phenotype. Expression levels of genes in normal tissue

(i.e., not undergoing prostate cancer) and in prostate cancer tissue (and in some cases, for varying severities of prostate cancer that relate to prognosis, as outlined below) are evaluated to provide expression profiles. An expression profile of a particular cell state or point of development is essentially a “fingerprint” of the state. While two states may have any particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is reflective of the state of the cell. By comparing expression profiles of cells in different states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained. Then, diagnosis may be performed or confirmed to determine whether a tissue sample has the gene expression profile of normal or cancerous tissue. This will provide for molecular diagnosis of related conditions.

“Differential expression,” or grammatical equivalents as used herein, refers to qualitative or quantitative differences in the temporal and/or cellular gene expression patterns within and among cells and tissue. Thus, a differentially expressed gene can qualitatively have its expression altered, including an activation or inactivation, in, e.g., normal versus prostate cancer tissue. Genes may be turned on or turned off in a particular state, relative to another state thus permitting comparison of two or more states. A qualitatively regulated gene will exhibit an expression pattern within a state or cell type which is detectable by standard techniques. Some genes will be expressed in one state or cell type, but not in both. Alternatively, the difference in expression may be quantitative, e.g., in that expression is increased or decreased; i.e., gene expression is either upregulated, resulting in an increased amount of transcript, or downregulated, resulting in a decreased amount of transcript. The degree to which expression differs need only be large enough to quantify via standard characterization techniques as outlined below, such as by use of Affymetrix GeneChip™ expression arrays, Lockhart, *Nature Biotechnology* 14:1675-1680 (1996), hereby expressly incorporated by reference. Other techniques include, but are not limited to, quantitative reverse transcriptase PCR, northern analysis and RNase protection. As outlined above, preferably the change in expression (i.e., upregulation or downregulation) is at least about 50%, more preferably at least about 100%, more preferably at least about 150%, more preferably at least about 200%, with from 300 to at least 1000% being especially preferred.

Evaluation may be at the gene transcript, or the protein level. The amount of gene expression may be monitored using nucleic acid probes to the DNA or RNA equivalent of the gene transcript, and the quantification of gene expression levels, or, alternatively, the final gene product itself (protein) can be monitored, e.g., with antibodies to the prostate cancer protein and standard immunoassays (ELISAs, etc.) or other techniques, including mass spectroscopy assays, 2D gel electrophoresis assays, etc. Proteins corresponding to prostate cancer genes, i.e., those identified as being important in a prostate cancer phenotype, can be evaluated in a prostate cancer diagnostic test.

In a preferred embodiment, gene expression monitoring is performed simultaneously on a number of genes. Multiple protein expression monitoring can be performed as well. Similarly, these assays may be performed on an individual basis as well.

In this embodiment, the prostate cancer nucleic acid probes are attached to biochips as outlined herein for the detection and quantification of prostate cancer sequences in a particular cell. The assays are further described below in the example. PCR techniques can be used to provide greater sensitivity.

In a preferred embodiment nucleic acids encoding the prostate cancer protein are detected. Although DNA or RNA encoding the prostate cancer protein may be detected, of particular interest are methods wherein an mRNA encoding a prostate cancer protein is detected. Probes to detect mRNA can be a nucleotide/deoxynucleotide probe that is complementary to and hybridizes with the mRNA and includes, but is not limited to, oligonucleotides, cDNA or RNA. Probes also should contain a detectable label, as defined herein. In one method the mRNA is detected after immobilizing the nucleic acid to be examined on a solid support such as nylon membranes and hybridizing the probe with the sample. Following washing to remove the non-specifically bound probe, the label is detected. In another method detection of the mRNA is performed in situ. In this method permeabilized cells or tissue samples are contacted with a detectably labeled nucleic acid probe for sufficient time to allow the probe to hybridize with the target mRNA. Following washing to remove the non-specifically bound probe, the label is detected. For example a digoxigenin labeled riboprobe (RNA probe) that is complementary to the mRNA encoding a prostate cancer protein is detected by binding the digoxigenin with an anti-digoxigenin

secondary antibody and developed with nitro blue tetrazolium and 5-bromo-4-chloro-3-indoyl phosphate.

In a preferred embodiment, various proteins from the three classes of proteins as described herein (secreted, transmembrane or intracellular proteins) are used in diagnostic assays. The prostate cancer proteins, antibodies, nucleic acids, modified proteins and cells containing prostate cancer sequences are used in diagnostic assays. This can be performed on an individual gene or corresponding polypeptide level. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes and/or corresponding polypeptides.

As described and defined herein, prostate cancer proteins, including intracellular, transmembrane or secreted proteins, find use as markers of prostate cancer. Detection of these proteins in putative prostate cancer tissue allows for detection or diagnosis of prostate cancer. In one embodiment, antibodies are used to detect prostate cancer proteins. A preferred method separates proteins from a sample by electrophoresis on a gel (typically a denaturing and reducing protein gel, but may be another type of gel, including isoelectric focusing gels and the like). Following separation of proteins, the prostate cancer protein is detected, e.g., by immunoblotting with antibodies raised against the prostate cancer protein. Methods of immunoblotting are well known to those of ordinary skill in the art.

In another preferred method, antibodies to the prostate cancer protein find use in *in situ* imaging techniques, e.g., in histology (e.g., *Methods in Cell Biology: Antibodies in Cell Biology*, volume 37 (Asai, ed. 1993)). In this method cells are contacted with from one to many antibodies to the prostate cancer protein(s). Following washing to remove non-specific antibody binding, the presence of the antibody or antibodies is detected. In one embodiment the antibody is detected by incubating with a secondary antibody that contains a detectable label. In another method the primary antibody to the prostate cancer protein(s) contains a detectable label, e.g. an enzyme marker that can act on a substrate. In another preferred embodiment each one of multiple primary antibodies contains a distinct and detectable label. This method finds particular use in simultaneous screening for a plurality of prostate cancer proteins. As will be appreciated by one of ordinary skill in the art, many other histological imaging techniques are also provided by the invention.

In a preferred embodiment the label is detected in a fluorometer which has the ability to detect and distinguish emissions of different wavelengths. In addition, a fluorescence activated cell sorter (FACS) can be used in the method.

In another preferred embodiment, antibodies find use in diagnosing prostate cancer from blood, serum, plasma, stool, and other samples. Such samples, therefore, are useful as samples to be probed or tested for the presence of prostate cancer proteins. Antibodies can be used to detect a prostate cancer protein by previously described immunoassay techniques including ELISA, immunoblotting (western blotting), immunoprecipitation, BIACORE technology and the like. Conversely, the presence of antibodies may indicate an immune response against an endogenous prostate cancer protein.

In a preferred embodiment, *in situ* hybridization of labeled prostate cancer nucleic acid probes to tissue arrays is done. For example, arrays of tissue samples, including prostate cancer tissue and/or normal tissue, are made. *In situ* hybridization (*see, e.g.,* Ausubel, *supra*) is then performed. When comparing the fingerprints between an individual and a standard, the skilled artisan can make a diagnosis, a prognosis, or a prediction based on the findings. It is further understood that the genes which indicate the diagnosis may differ from those which indicate the prognosis and molecular profiling of the condition of the cells may lead to distinctions between responsive or refractory conditions or may be predictive of outcomes.

In a preferred embodiment, the prostate cancer proteins, antibodies, nucleic acids, modified proteins and cells containing prostate cancer sequences are used in prognosis assays. As above, gene expression profiles can be generated that correlate to prostate cancer, in terms of long term prognosis. Again, this may be done on either a protein or gene level, with the use of genes being preferred. As above, prostate cancer probes may be attached to biochips for the detection and quantification of prostate cancer sequences in a tissue or patient. The assays proceed as outlined above for diagnosis. PCR method may provide more sensitive and accurate quantification.

Assays for therapeutic compounds

In a preferred embodiment members of the proteins, nucleic acids, and antibodies as described herein are used in drug screening assays. The prostate cancer

proteins, antibodies, nucleic acids, modified proteins and cells containing prostate cancer sequences are used in drug screening assays or by evaluating the effect of drug candidates on a “gene expression profile” or expression profile of polypeptides. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent (e.g., Zlokarnik, *et al.*, *Science* 279:84-8 (1998); Heid, *Genome Res* 6:986-94, 1996).

In a preferred embodiment, the prostate cancer proteins, antibodies, nucleic acids, modified proteins and cells containing the native or modified prostate cancer proteins are used in screening assays. That is, the present invention provides novel methods for screening for compositions which modulate the prostate cancer phenotype or an identified physiological function of a prostate cancer protein. As above, this can be done on an individual gene level or by evaluating the effect of drug candidates on a “gene expression profile”. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent, see Zlokarnik, *supra*.

Having identified the differentially expressed genes herein, a variety of assays may be executed. In a preferred embodiment, assays may be run on an individual gene or protein level. That is, having identified a particular gene as up regulated in prostate cancer, test compounds can be screened for the ability to modulate gene expression or for binding to the prostate cancer protein. “Modulation” thus includes both an increase and a decrease in gene expression. The preferred amount of modulation will depend on the original change of the gene expression in normal versus tissue undergoing prostate cancer, with changes of at least 10%, preferably 50%, more preferably 100-300%, and in some embodiments 300-1000% or greater. Thus, if a gene exhibits a 4-fold increase in prostate cancer tissue compared to normal tissue, a decrease of about four-fold is often desired; similarly, a 10-fold decrease in prostate cancer tissue compared to normal tissue often provides a target value of a 10-fold increase in expression to be induced by the test compound.

The amount of gene expression may be monitored using nucleic acid probes and the quantification of gene expression levels, or, alternatively, the gene product itself can be monitored, e.g., through the use of antibodies to the prostate cancer protein and standard

immunoassays. Proteomics and separation techniques may also allow quantification of expression.

In a preferred embodiment, gene expression or protein monitoring of a number of entities, i.e., an expression profile, is monitored simultaneously. Such profiles will typically involve a plurality of those entities described herein..

In this embodiment, the prostate cancer nucleic acid probes are attached to biochips as outlined herein for the detection and quantification of prostate cancer sequences in a particular cell. Alternatively, PCR may be used. Thus, a series, e.g., of microtiter plate, may be used with dispensed primers in desired wells. A PCR reaction can then be performed and analyzed for each well.

Expression monitoring can be performed to identify compounds that modify the expression of one or more prostate cancer-associated sequences, e.g., a polynucleotide sequence set out in Tables 1-16. Generally, in a preferred embodiment, a test modulator is added to the cells prior to analysis. Moreover, screens are also provided to identify agents that modulate prostate cancer, modulate prostate cancer proteins, bind to a prostate cancer protein, or interfere with the binding of a prostate cancer protein and an antibody or other binding partner.

The term "test compound" or "drug candidate" or "modulator" or grammatical equivalents as used herein describes any molecule, e.g., protein, oligopeptide, small organic molecule, polysaccharide, polynucleotide, etc., to be tested for the capacity to directly or indirectly alter the prostate cancer phenotype or the expression of a prostate cancer sequence, e.g., a nucleic acid or protein sequence. In preferred embodiments, modulators alter expression profiles, or expression profile nucleic acids or proteins provided herein. In one embodiment, the modulator suppresses a prostate cancer phenotype, e.g. to a normal tissue fingerprint. In another embodiment, a modulator induced a prostate cancer phenotype. Generally, a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a differential response to the various concentrations. Typically, one of these concentrations serves as a negative control, i.e., at zero concentration or below the level of detection.

Drug candidates encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of

more than 100 and less than about 2,500 daltons. Preferred small molecules are less than 2000, or less than 1500 or less than 1000 or less than 500 D. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs or combinations thereof. Particularly preferred are peptides.

In one aspect, a modulator will neutralize the effect of a prostate cancer protein. By “neutralize” is meant that activity of a protein is inhibited or blocked and the consequent effect on the cell.

In certain embodiments, combinatorial libraries of potential modulators will be screened for an ability to bind to a prostate cancer polypeptide or to modulate activity. Conventionally, new chemical entities with useful properties are generated by identifying a chemical compound (called a “lead compound”) with some desirable property or activity, e.g., inhibiting activity, creating variants of the lead compound, and evaluating the property and activity of those variant compounds. Often, high throughput screening (HTS) methods are employed for such an analysis.

In one preferred embodiment, high throughput screening methods involve providing a library containing a large number of potential therapeutic compounds (candidate compounds). Such “combinatorial chemical libraries” are then screened in one or more assays to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional “lead compounds” or can themselves be used as potential or actual therapeutics.

A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis by combining a number of chemical “building blocks” such as reagents. For example, a linear combinatorial chemical library, such as a polypeptide (e.g., mutein) library, is formed by combining a set of chemical building blocks called amino acids in every possible way for a given compound

length (i.e., the number of amino acids in a polypeptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks (Gallop *et al.*, *J. Med. Chem.* 37(9):1233-1251 (1994)).

Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (*see, e.g.*, U.S. Patent No. 5,010,175, Furka, *Pept. Prot. Res.* 37:487-493 (1991), Houghton *et al.*, *Nature*, 354:84-88 (1991)), peptoids (PCT Publication No WO 91/19735), encoded peptides (PCT Publication WO 93/20242), random bio-oligomers (PCT Publication WO 92/00091), benzodiazepines (U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs *et al.*, *Proc. Nat. Acad. Sci. USA* 90:6909-6913 (1993)), vinyllogous polypeptides (Hagihara *et al.*, *J. Amer. Chem. Soc.* 114:6568 (1992)), nonpeptidal peptidomimetics with a Beta-D-Glucose scaffolding (Hirschmann *et al.*, *J. Amer. Chem. Soc.* 114:9217-9218 (1992)), analogous organic syntheses of small compound libraries (Chen *et al.*, *J. Amer. Chem. Soc.* 116:2661 (1994)), oligocarbamates (Cho, *et al.*, *Science* 261:1303 (1993)), and/or peptidyl phosphonates (Campbell *et al.*, *J. Org. Chem.* 59:658 (1994)). *See, generally*, Gordon *et al.*, *J. Med. Chem.* 37:1385 (1994), nucleic acid libraries (*see, e.g.*, Strategene, Corp.), peptide nucleic acid libraries (*see, e.g.*, U.S. Patent 5,539,083), antibody libraries (*see, e.g.*, Vaughn *et al.*, *Nature Biotechnology* 14(3):309-314 (1996), and PCT/US96/10287), carbohydrate libraries (*see, e.g.*, Liang *et al.*, *Science* 274:1520-1522 (1996), and U.S. Patent No. 5,593,853), and small organic molecule libraries (*see, e.g.*, benzodiazepines, Baum, C&EN, Jan 18, page 33 (1993); isoprenoids, U.S. Patent No. 5,569,588; thiazolidinones and metathiazanones, U.S. Patent No. 5,549,974; pyrrolidines, U.S. Patent Nos. 5,525,735 and 5,519,134; morpholino compounds, U.S. Patent No. 5,506,337; benzodiazepines, U.S. Patent No. 5,288,514; and the like).

Devices for the preparation of combinatorial libraries are commercially available (*see, e.g.*, 357 MPS, 390 MPS, Advanced Chem Tech, Louisville KY, Symphony, Rainin, Woburn, MA, 433A Applied Biosystems, Foster City, CA, 9050 Plus, Millipore, Bedford, MA).

A number of well known robotic systems have also been developed for solution phase chemistries. These systems include automated workstations like the automated synthesis apparatus developed by Takeda Chemical Industries, LTD. (Osaka,

Japan) and many robotic systems utilizing robotic arms (Zymate II, Zymark Corporation, Hopkinton, Mass.; Orca, Hewlett-Packard, Palo Alto, Calif.), which mimic the manual synthetic operations performed by a chemist. Any of the above devices are suitable for use with the present invention. The nature and implementation of modifications to these devices (if any) so that they can operate as discussed herein will be apparent to persons skilled in the relevant art. In addition, numerous combinatorial libraries are themselves commercially available (*see, e.g.*, ComGenex, Princeton, N.J., Asinex, Moscow, Ru, Tripos, Inc., St. Louis, MO, ChemStar, Ltd, Moscow, RU, 3D Pharmaceuticals, Exton, PA, Martek Biosciences, Columbia, MD, *etc.*).

The assays to identify modulators are amenable to high throughput screening. Preferred assays thus detect enhancement or inhibition of prostate cancer gene transcription, inhibition or enhancement of polypeptide expression, and inhibition or enhancement of polypeptide activity.

High throughput assays for the presence, absence, quantification, or other properties of particular nucleic acids or protein products are well known to those of skill in the art. Similarly, binding assays and reporter gene assays are similarly well known. Thus, *e.g.*, U.S. Patent No. 5,559,410 discloses high throughput screening methods for proteins, U.S. Patent No. 5,585,639 discloses high throughput screening methods for nucleic acid binding (*i.e.*, in arrays), while U.S. Patent Nos. 5,576,220 and 5,541,061 disclose high throughput methods of screening for ligand/antibody binding.

In addition, high throughput screening systems are commercially available (*see, e.g.*, Zymark Corp., Hopkinton, MA; Air Technical Industries, Mentor, OH; Beckman Instruments, Inc. Fullerton, CA; Precision Systems, Inc., Natick, MA, *etc.*). These systems typically automate entire procedures, including all sample and reagent pipetting, liquid dispensing, timed incubations, and final readings of the microplate in detector(s) appropriate for the assay. These configurable systems provide high throughput and rapid start up as well as a high degree of flexibility and customization. The manufacturers of such systems provide detailed protocols for various high throughput systems. Thus, *e.g.*, Zymark Corp. provides technical bulletins describing screening systems for detecting the modulation of gene transcription, ligand binding, and the like.

In one embodiment, modulators are proteins, often naturally occurring proteins or fragments of naturally occurring proteins. Thus, *e.g.*, cellular extracts containing proteins, or random or directed digests of proteinaceous cellular extracts, may be used. In this way libraries of proteins may be made for screening in the methods of the invention.

5 Particularly preferred in this embodiment are libraries of bacterial, fungal, viral, and mammalian proteins, with the latter being preferred, and human proteins being especially preferred. Particularly useful test compound will be directed to the class of proteins to which the target belongs, *e.g.*, substrates for enzymes or ligands and receptors.

In a preferred embodiment, modulators are peptides of from about 5 to about
10 30 amino acids, with from about 5 to about 20 amino acids being preferred, and from about 7 to about 15 being particularly preferred. The peptides may be digests of naturally occurring proteins as is outlined above, random peptides, or “biased” random peptides. By “randomized” or grammatical equivalents herein is meant that each nucleic acid and peptide consists of essentially random nucleotides and amino acids, respectively. Since generally
15 these random peptides (or nucleic acids, discussed below) are chemically synthesized, they may incorporate any nucleotide or amino acid at any position. The synthetic process can be designed to generate randomized proteins or nucleic acids, to allow the formation of all or most of the possible combinations over the length of the sequence, thus forming a library of randomized candidate bioactive proteinaceous agents.

20 In one embodiment, the library is fully randomized, with no sequence preferences or constants at any position. In a preferred embodiment, the library is biased. That is, some positions within the sequence are either held constant, or are selected from a limited number of possibilities. For example, in a preferred embodiment, the nucleotides or amino acid residues are randomized within a defined class, *e.g.*, of hydrophobic amino acids,
25 hydrophilic residues, sterically biased (either small or large) residues, towards the creation of nucleic acid binding domains, the creation of cysteines, for cross-linking, prolines for SH-3 domains, serines, threonines, tyrosines or histidines for phosphorylation sites, etc., or to purines, etc.

Modulators of prostate cancer can also be nucleic acids, as defined below. As
30 described above generally for proteins, nucleic acid modulating agents may be naturally occurring nucleic acids, random nucleic acids, or “biased” random nucleic acids. For

example, digests of procaryotic or eucaryotic genomes may be used as is outlined above for proteins.

In certain embodiments, the activity of a prostate cancer-associated protein is down-regulated, or entirely inhibited, by the use of antisense polynucleotide, *i.e.*, a nucleic acid complementary to, and which can preferably hybridize specifically to, a coding mRNA nucleic acid sequence, e.g., a prostate cancer protein mRNA, or a subsequence thereof. Binding of the antisense polynucleotide to the mRNA reduces the translation and/or stability of the mRNA.

In the context of this invention, antisense polynucleotides can comprise naturally-occurring nucleotides, or synthetic species formed from naturally-occurring subunits or their close homologs. Antisense polynucleotides may also have altered sugar moieties or inter-sugar linkages. Exemplary among these are the phosphorothioate and other sulfur containing species which are known for use in the art. Analogs are comprehended by this invention so long as they function effectively to hybridize with the prostate cancer protein mRNA. *See, e.g.*, Isis Pharmaceuticals, Carlsbad, CA; Sequitor, Inc., Natick, MA.

Such antisense polynucleotides can readily be synthesized using recombinant means, or can be synthesized *in vitro*. Equipment for such synthesis is sold by several vendors, including Applied Biosystems. The preparation of other oligonucleotides such as phosphorothioates and alkylated derivatives is also well known to those of skill in the art.

Antisense molecules as used herein include antisense or sense oligonucleotides. Sense oligonucleotides can, e.g., be employed to block transcription by binding to the anti-sense strand. The antisense and sense oligonucleotide comprise a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target mRNA (sense) or DNA (antisense) sequences for prostate cancer molecules. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment generally at least about 14 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, e.g., Stein & Cohen (*Cancer Res.* 48:2659 (1988) and van der Krol *et al.* (*BioTechniques* 6:958 (1988)).

In addition to antisense polynucleotides, ribozymes can be used to target and inhibit transcription of prostate cancer-associated nucleotide sequences. A ribozyme is an

RNA molecule that catalytically cleaves other RNA molecules. Different kinds of ribozymes have been described, including group I ribozymes, hammerhead ribozymes, hairpin ribozymes, RNase P, and axhead ribozymes (*see, e.g., Castanotto et al., Adv. in Pharmacology* 25: 289-317 (1994) for a general review of the properties of different ribozymes).

The general features of hairpin ribozymes are described, e.g., in Hampel *et al., Nucl. Acids Res.* 18:299-304 (1990); European Patent Publication No. 0 360 257; U.S. Patent No. 5,254,678. Methods of preparing are well known to those of skill in the art (*see, e.g., WO 94/26877; Ojwang et al., Proc. Natl. Acad. Sci. USA* 90:6340-6344 (1993); Yamada *et al., Human Gene Therapy* 1:39-45 (1994); Leavitt *et al., Proc. Natl. Acad. Sci. USA* 92:699-703 (1995); Leavitt *et al., Human Gene Therapy* 5:1151-120 (1994); and Yamada *et al., Virology* 205: 121-126 (1994)).

Polynucleotide modulators of prostate cancer may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell. Alternatively, a polynucleotide modulator of prostate cancer may be introduced into a cell containing the target nucleic acid sequence, e.g., by formation of an polynucleotide-lipid complex, as described in WO 90/10448. It is understood that the use of antisense molecules or knock out and knock in models may also be used in screening assays as discussed above, in addition to methods of treatment.

As noted above, gene expression monitoring is conveniently used to test candidate modulators (*e.g., protein, nucleic acid or small molecule*). After the candidate agent has been added and the cells allowed to incubate for some period of time, the sample containing a target sequence to be analyzed is added to the biochip. If required, the target sequence is prepared using known techniques. For example, the sample may be treated to lyse the cells, using known lysis buffers, electroporation, etc., with purification and/or amplification such as PCR performed as appropriate. For example, an *in vitro* transcription

with labels covalently attached to the nucleotides is performed. Generally, the nucleic acids are labeled with biotin-FITC or PE, or with cy3 or cy5.

In a preferred embodiment, the target sequence is labeled with, e.g., a fluorescent, a chemiluminescent, a chemical, or a radioactive signal, to provide a means of detecting the target sequence's specific binding to a probe. The label also can be an enzyme, such as, alkaline phosphatase or horseradish peroxidase, which when provided with an appropriate substrate produces a product that can be detected. Alternatively, the label can be a labeled compound or small molecule, such as an enzyme inhibitor, that binds but is not catalyzed or altered by the enzyme. The label also can be a moiety or compound, such as, an epitope tag or biotin which specifically binds to streptavidin. For the example of biotin, the streptavidin is labeled as described above, thereby, providing a detectable signal for the bound target sequence. Unbound labeled streptavidin is typically removed prior to analysis.

As will be appreciated by those in the art, these assays can be direct hybridization assays or can comprise "sandwich assays", which include the use of multiple probes, as is generally outlined in U.S. Patent Nos. 5,681,702, 5,597,909, 5,545,730, 5,594,117, 5,591,584, 5,571,670, 5,580,731, 5,571,670, 5,591,584, 5,624,802, 5,635,352, 5,594,118, 5,359,100, 5,124,246 and 5,681,697, all of which are hereby incorporated by reference. In this embodiment, in general, the target nucleic acid is prepared as outlined above, and then added to the biochip comprising a plurality of nucleic acid probes, under conditions that allow the formation of a hybridization complex.

A variety of hybridization conditions may be used in the present invention, including high, moderate and low stringency conditions as outlined above. The assays are generally run under stringency conditions which allows formation of the label probe hybridization complex only in the presence of target. Stringency can be controlled by altering a step parameter that is a thermodynamic variable, including, but not limited to, temperature, formamide concentration, salt concentration, chaotropic salt concentration pH, organic solvent concentration, etc.

These parameters may also be used to control non-specific binding, as is generally outlined in U.S. Patent No. 5,681,697. Thus it may be desirable to perform certain steps at higher stringency conditions to reduce non-specific binding.

The reactions outlined herein may be accomplished in a variety of ways. Components of the reaction may be added simultaneously, or sequentially, in different orders, with preferred embodiments outlined below. In addition, the reaction may include a variety of other reagents. These include salts, buffers, neutral proteins, e.g. albumin, detergents, *etc.* which may be used to facilitate optimal hybridization and detection, and/or reduce non-specific or background interactions. Reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, *etc.*, may also be used as appropriate, depending on the sample preparation methods and purity of the target.

The assay data are analyzed to determine the expression levels, and changes in expression levels as between states, of individual genes, forming a gene expression profile.

Screens are performed to identify modulators of the prostate cancer phenotype. In one embodiment, screening is performed to identify modulators that can induce or suppress a particular expression profile, thus preferably generating the associated phenotype. In another embodiment, *e.g.*, for diagnostic applications, having identified differentially expressed genes important in a particular state, screens can be performed to identify modulators that alter expression of individual genes. In another embodiment, screening is performed to identify modulators that alter a biological function of the expression product of a differentially expressed gene. Again, having identified the importance of a gene in a particular state, screens are performed to identify agents that bind and/or modulate the biological activity of the gene product.

In addition screens can be done for genes that are induced in response to a candidate agent. After identifying a modulator based upon its ability to suppress a prostate cancer expression pattern leading to a normal expression pattern, or to modulate a single prostate cancer gene expression profile so as to mimic the expression of the gene from normal tissue, a screen as described above can be performed to identify genes that are specifically modulated in response to the agent. Comparing expression profiles between normal tissue and agent treated prostate cancer tissue reveals genes that are not expressed in normal tissue or prostate cancer tissue, but are expressed in agent treated tissue. These agent-specific sequences can be identified and used by methods described herein for prostate cancer genes or proteins. In particular these sequences and the proteins they encode find use in marking or identifying agent treated cells. In addition, antibodies can be raised against the

agent induced proteins and used to target novel therapeutics to the treated prostate cancer tissue sample.

Thus, in one embodiment, a test compound is administered to a population of prostate cancer cells, that have an associated prostate cancer expression profile. By “administration” or “contacting” herein is meant that the candidate agent is added to the cells in such a manner as to allow the agent to act upon the cell, whether by uptake and intracellular action, or by action at the cell surface. In some embodiments, nucleic acid encoding a proteinaceous candidate agent (i.e., a peptide) may be put into a viral construct such as an adenoviral or retroviral construct, and added to the cell, such that expression of the peptide agent is accomplished, e.g., PCT US97/01019. Regulatable gene therapy systems can also be used.

Once the test compound has been administered to the cells, the cells can be washed if desired and are allowed to incubate under preferably physiological conditions for some period of time. The cells are then harvested and a new gene expression profile is generated, as outlined herein.

Thus, e.g., prostate cancer tissue may be screened for agents that modulate, e.g., induce or suppress the prostate cancer phenotype. A change in at least one gene, preferably many, of the expression profile indicates that the agent has an effect on prostate cancer activity. By defining such a signature for the prostate cancer phenotype, screens for new drugs that alter the phenotype can be devised. With this approach, the drug target need not be known and need not be represented in the original expression screening platform, nor does the level of transcript for the target protein need to change.

In a preferred embodiment, as outlined above, screens may be done on individual genes and gene products (proteins). That is, having identified a particular differentially expressed gene as important in a particular state, screening of modulators of either the expression of the gene or the gene product itself can be done. The gene products of differentially expressed genes are sometimes referred to herein as “prostate cancer proteins” or a “prostate cancer modulatory protein”. The prostate cancer modulatory protein may be a fragment, or alternatively, be the full length protein to the fragment encoded by the nucleic acids of Tables 1-16. Preferably, the prostate cancer modulatory protein is a fragment. In a preferred embodiment, the prostate cancer amino acid sequence which is used to determine

sequence identity or similarity is encoded by a nucleic acid of Tables 1-16. In another embodiment, the sequences are naturally occurring allelic variants of a protein encoded by a nucleic acid of Tables 1-16. In another embodiment, the sequences are sequence variants as further described herein.

5 Preferably, the prostate cancer modulatory protein is a fragment of approximately 14 to 24 amino acids long. More preferably the fragment is a soluble fragment. Preferably, the fragment includes a non-transmembrane region. In a preferred embodiment, the fragment has an N-terminal Cys to aid in solubility. In one embodiment, the C-terminus of the fragment is kept as a free acid and the N-terminus is a free amine to aid in
10 coupling, i.e., to cysteine.

 In one embodiment the prostate cancer proteins are conjugated to an immunogenic agent as discussed herein. In one embodiment the prostate cancer protein is conjugated to BSA.

 Measurements of prostate cancer polypeptide activity, or of prostate cancer or
15 the prostate cancer phenotype can be performed using a variety of assays. For example, the effects of the test compounds upon the function of the prostate cancer polypeptides can be measured by examining parameters described above. A suitable physiological change that affects activity can be used to assess the influence of a test compound on the polypeptides of this invention. When the functional consequences are determined using intact cells or
20 animals, one can also measure a variety of effects such as, in the case of prostate cancer associated with tumors, tumor growth, tumor metastasis, neovascularization, hormone release, transcriptional changes to both known and uncharacterized genetic markers (e.g., northern blots), changes in cell metabolism such as cell growth or pH changes, and changes in intracellular second messengers such as cGMP. In the assays of the invention, mammalian
25 prostate cancer polypeptide is typically used, e.g., mouse, preferably human.

 Assays to identify compounds with modulating activity can be performed *in vitro*. For example, a prostate cancer polypeptide is first contacted with a potential modulator and incubated for a suitable amount of time, e.g., from 0.5 to 48 hours. In one embodiment, the prostate cancer polypeptide levels are determined *in vitro* by measuring the level of
30 protein or mRNA. The level of protein is measured using immunoassays such as western blotting, ELISA and the like with an antibody that selectively binds to the prostate cancer

polypeptide or a fragment thereof. For measurement of mRNA, amplification, e.g., using PCR, LCR, or hybridization assays, e.g., northern hybridization, RNase protection, dot blotting, are preferred. The level of protein or mRNA is detected using directly or indirectly labeled detection agents, e.g., fluorescently or radioactively labeled nucleic acids,
5 radioactively or enzymatically labeled antibodies, and the like, as described herein.

Alternatively, a reporter gene system can be devised using the prostate cancer protein promoter operably linked to a reporter gene such as luciferase, green fluorescent protein, CAT, or β -gal. The reporter construct is typically transfected into a cell. After treatment with a potential modulator, the amount of reporter gene transcription, translation, or
10 activity is measured according to standard techniques known to those of skill in the art.

In a preferred embodiment, as outlined above, screens may be done on individual genes and gene products (proteins). That is, having identified a particular differentially expressed gene as important in a particular state, screening of modulators of the expression of the gene or the gene product itself can be done. The gene products of
15 differentially expressed genes are sometimes referred to herein as "prostate cancer proteins." The prostate cancer protein may be a fragment, or alternatively, be the full length protein to a fragment shown herein.

In one embodiment, screening for modulators of expression of specific genes is performed. Typically, the expression of only one or a few genes are evaluated. In another
20 embodiment, screens are designed to first find compounds that bind to differentially expressed proteins. These compounds are then evaluated for the ability to modulate differentially expressed activity. Moreover, once initial candidate compounds are identified, variants can be further screened to better evaluate structure activity relationships.

In a preferred embodiment, binding assays are done. In general, purified or
25 isolated gene product is used; that is, the gene products of one or more differentially expressed nucleic acids are made. For example, antibodies are generated to the protein gene products, and standard immunoassays are run to determine the amount of protein present. Alternatively, cells comprising the prostate cancer proteins can be used in the assays.

Thus, in a preferred embodiment, the methods comprise combining a prostate
30 cancer protein and a candidate compound, and determining the binding of the compound to the prostate cancer protein. Preferred embodiments utilize the human prostate cancer protein,

although other mammalian proteins may also be used, e.g. for the development of animal models of human disease. In some embodiments, as outlined herein, variant or derivative prostate cancer proteins may be used.

Generally, in a preferred embodiment of the methods herein, the prostate cancer protein or the candidate agent is non-diffusably bound to an insoluble support having isolated sample receiving areas (e.g. a microtiter plate, an array, etc.). The insoluble supports may be made of any composition to which the compositions can be bound, is readily separated from soluble material, and is otherwise compatible with the overall method of screening. The surface of such supports may be solid or porous and of any convenient shape. Examples of suitable insoluble supports include microtiter plates, arrays, membranes and beads. These are typically made of glass, plastic (e.g., polystyrene), polysaccharides, nylon or nitrocellulose, teflon™, etc. Microtiter plates and arrays are especially convenient because a large number of assays can be carried out simultaneously, using small amounts of reagents and samples. The particular manner of binding of the composition is not crucial so long as it is compatible with the reagents and overall methods of the invention, maintains the activity of the composition and is nondiffusable. Preferred methods of binding include the use of antibodies (which do not sterically block either the ligand binding site or activation sequence when the protein is bound to the support), direct binding to “sticky” or ionic supports, chemical crosslinking, the synthesis of the protein or agent on the surface, etc. Following binding of the protein or agent, excess unbound material is removed by washing. The sample receiving areas may then be blocked through incubation with bovine serum albumin (BSA), casein or other innocuous protein or other moiety.

In a preferred embodiment, the prostate cancer protein is bound to the support, and a test compound is added to the assay. Alternatively, the candidate agent is bound to the support and the prostate cancer protein is added. Novel binding agents include specific antibodies, non-natural binding agents identified in screens of chemical libraries, peptide analogs, etc. Of particular interest are screening assays for agents that have a low toxicity for human cells. A wide variety of assays may be used for this purpose, including labeled in vitro protein-protein binding assays, electrophoretic mobility shift assays, immunoassays for protein binding, functional assays (phosphorylation assays, etc.) and the like.

The determination of the binding of the test modulating compound to the prostate cancer protein may be done in a number of ways. In a preferred embodiment, the compound is labeled, and binding determined directly, e.g., by attaching all or a portion of the prostate cancer protein to a solid support, adding a labeled candidate agent (e.g., a
5 fluorescent label), washing off excess reagent, and determining whether the label is present on the solid support. Various blocking and washing steps may be utilized as appropriate.

In some embodiments, only one of the components is labeled, e.g., the proteins (or proteinaceous candidate compounds) can be labeled. Alternatively, more than one component can be labeled with different labels, e.g., ^{125}I for the proteins and a fluorophor
10 for the compound. Proximity reagents, e.g., quenching or energy transfer reagents are also useful.

In one embodiment, the binding of the test compound is determined by competitive binding assay. The competitor is a binding moiety known to bind to the target molecule (i.e., a prostate cancer protein), such as an antibody, peptide, binding partner,
15 ligand, etc. Under certain circumstances, there may be competitive binding between the compound and the binding moiety, with the binding moiety displacing the compound. In one embodiment, the test compound is labeled. Either the compound, or the competitor, or both, is added first to the protein for a time sufficient to allow binding, if present. Incubations may be performed at a temperature which facilitates optimal activity, typically between 4 and
20 40°C. Incubation periods are typically optimized, e.g., to facilitate rapid high throughput screening. Typically between 0.1 and 1 hour will be sufficient. Excess reagent is generally removed or washed away. The second component is then added, and the presence or absence of the labeled component is followed, to indicate binding.

In a preferred embodiment, the competitor is added first, followed by the test
25 compound. Displacement of the competitor is an indication that the test compound is binding to the prostate cancer protein and thus is capable of binding to, and potentially modulating, the activity of the prostate cancer protein. In this embodiment, either component can be labeled. Thus, e.g., if the competitor is labeled, the presence of label in the wash solution indicates displacement by the agent. Alternatively, if the test compound is labeled, the
30 presence of the label on the support indicates displacement.

In an alternative embodiment, the test compound is added first, with incubation and washing, followed by the competitor. The absence of binding by the competitor may indicate that the test compound is bound to the prostate cancer protein with a higher affinity. Thus, if the test compound is labeled, the presence of the label on the support, coupled with a lack of competitor binding, may indicate that the test compound is capable of binding to the prostate cancer protein.

In a preferred embodiment, the methods comprise differential screening to identify agents that are capable of modulating the activity of the prostate cancer proteins. In this embodiment, the methods comprise combining a prostate cancer protein and a competitor in a first sample. A second sample comprises a test compound, a prostate cancer protein, and a competitor. The binding of the competitor is determined for both samples, and a change, or difference in binding between the two samples indicates the presence of an agent capable of binding to the prostate cancer protein and potentially modulating its activity. That is, if the binding of the competitor is different in the second sample relative to the first sample, the agent is capable of binding to the prostate cancer protein.

Alternatively, differential screening is used to identify drug candidates that bind to the native prostate cancer protein, but cannot bind to modified prostate cancer proteins. The structure of the prostate cancer protein may be modeled, and used in rational drug design to synthesize agents that interact with that site. Drug candidates that affect the activity of a prostate cancer protein are also identified by screening drugs for the ability to either enhance or reduce the activity of the protein.

Positive controls and negative controls may be used in the assays. Preferably control and test samples are performed in at least triplicate to obtain statistically significant results. Incubation of all samples is for a time sufficient for the binding of the agent to the protein. Following incubation, samples are washed free of non-specifically bound material and the amount of bound, generally labeled agent determined. For example, where a radiolabel is employed, the samples may be counted in a scintillation counter to determine the amount of bound compound.

A variety of other reagents may be included in the screening assays. These include reagents like salts, neutral proteins, e.g. albumin, detergents, etc. which may be used to facilitate optimal protein-protein binding and/or reduce non-specific or background

interactions. Also reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used. The mixture of components may be added in an order that provides for the requisite binding.

In a preferred embodiment, the invention provides methods for screening for a compound capable of modulating the activity of a prostate cancer protein. The methods comprise adding a test compound, as defined above, to a cell comprising prostate cancer proteins. Preferred cell types include almost any cell. The cells contain a recombinant nucleic acid that encodes a prostate cancer protein. In a preferred embodiment, a library of candidate agents are tested on a plurality of cells.

In one aspect, the assays are evaluated in the presence or absence or previous or subsequent exposure of physiological signals, e.g. hormones, antibodies, peptides, antigens, cytokines, growth factors, action potentials, pharmacological agents including chemotherapeutics, radiation, carcinogenics, or other cells (i.e. cell-cell contacts). In another example, the determinations are determined at different stages of the cell cycle process.

In this way, compounds that modulate prostate cancer agents are identified. Compounds with pharmacological activity are able to enhance or interfere with the activity of the prostate cancer protein. Once identified, similar structures are evaluated to identify critical structural feature of the compound.

In one embodiment, a method of inhibiting prostate cancer cell division is provided. The method comprises administration of a prostate cancer inhibitor. In another embodiment, a method of inhibiting prostate cancer is provided. The method comprises administration of a prostate cancer inhibitor. In a further embodiment, methods of treating cells or individuals with prostate cancer are provided. The method comprises administration of a prostate cancer inhibitor.

In one embodiment, a prostate cancer inhibitor is an antibody as discussed above. In another embodiment, the prostate cancer inhibitor is an antisense molecule.

A variety of cell growth, proliferation, and metastasis assays are known to those of skill in the art, as described below.

Soft agar growth or colony formation in suspension

Normal cells require a solid substrate to attach and grow. When the cells are transformed, they lose this phenotype and grow detached from the substrate. For example,

transformed cells can grow in stirred suspension culture or suspended in semi-solid media, such as semi-solid or soft agar. The transformed cells, when transfected with tumor suppressor genes, regenerate normal phenotype and require a solid substrate to attach and grow. Soft agar growth or colony formation in suspension assays can be used to identify
5 modulators of prostate cancer sequences, which when expressed in host cells, inhibit abnormal cellular proliferation and transformation. A therapeutic compound would reduce or eliminate the host cells' ability to grow in stirred suspension culture or suspended in semi-solid media, such as semi-solid or soft.

Techniques for soft agar growth or colony formation in suspension assays are
10 described in Freshney, *Culture of Animal Cells a Manual of Basic Technique* (3rd ed., 1994), herein incorporated by reference. *See also*, the methods section of Garkavtsev *et al.* (1996), *supra*, herein incorporated by reference.

Contact inhibition and density limitation of growth

Normal cells typically grow in a flat and organized pattern in a petri dish until
15 they touch other cells. When the cells touch one another, they are contact inhibited and stop growing. When cells are transformed, however, the cells are not contact inhibited and continue to grow to high densities in disorganized foci. Thus, the transformed cells grow to a higher saturation density than normal cells. This can be detected morphologically by the formation of a disoriented monolayer of cells or rounded cells in foci within the regular
20 pattern of normal surrounding cells. Alternatively, labeling index with (³H)-thymidine at saturation density can be used to measure density limitation of growth. *See* Freshney (1994), *supra*. The transformed cells, when transfected with tumor suppressor genes, regenerate a normal phenotype and become contact inhibited and would grow to a lower density.

In this assay, labeling index with (³H)-thymidine at saturation density is a
25 preferred method of measuring density limitation of growth. Transformed host cells are transfected with a prostate cancer-associated sequence and are grown for 24 hours at saturation density in non-limiting medium conditions. The percentage of cells labeling with (³H)-thymidine is determined autoradiographically. *See*, Freshney (1994), *supra*.

Growth factor or serum dependence

Transformed cells have a lower serum dependence than their normal counterparts (*see, e.g.,* Temin, *J. Natl. Cancer Inst.* 37:167-175 (1966); Eagle *et al.*, *J. Exp. Med.* 131:836-879 (1970)); Freshney, *supra*. This is in part due to release of various growth factors by the transformed cells. Growth factor or serum dependence of transformed host cells can be compared with that of control.

Tumor specific markers levels

Tumor cells release an increased amount of certain factors (hereinafter “tumor specific markers”) than their normal counterparts. For example, plasminogen activator (PA) is released from human glioma at a higher level than from normal brain cells (*see, e.g.,* Gullino, *Angiogenesis, tumor vascularization, and potential interference with tumor growth*. in *Biological Responses in Cancer*, pp. 178-184 (Mihich (ed.) 1985)). Similarly, Tumor angiogenesis factor (TAF) is released at a higher level in tumor cells than their normal counterparts. *See, e.g.,* Folkman, *Angiogenesis and Cancer, Sem Cancer Biol.* (1992)).

Various techniques which measure the release of these factors are described in Freshney (1994), *supra*. Also, *see, Unkless et al.*, *J. Biol. Chem.* 249:4295-4305 (1974); Strickland & Beers, *J. Biol. Chem.* 251:5694-5702 (1976); Whur *et al.*, *Br. J. Cancer* 42:305-312 (1980); Gullino, *Angiogenesis, tumor vascularization, and potential interference with tumor growth*. in *Biological Responses in Cancer*, pp. 178-184 (Mihich (ed.) 1985); Freshney *Anticancer Res.* 5:111-130 (1985).

Invasiveness into Matrigel

The degree of invasiveness into Matrigel or some other extracellular matrix constituent can be used as an assay to identify compounds that modulate prostate cancer-associated sequences. Tumor cells exhibit a good correlation between malignancy and invasiveness of cells into Matrigel or some other extracellular matrix constituent. In this assay, tumorigenic cells are typically used as host cells. Expression of a tumor suppressor gene in these host cells would decrease invasiveness of the host cells.

Techniques described in Freshney (1994), *supra*, can be used. Briefly, the level of invasion of host cells can be measured by using filters coated with Matrigel or some

other extracellular matrix constituent. Penetration into the gel, or through to the distal side of the filter, is rated as invasiveness, and rated histologically by number of cells and distance moved, or by prelabeling the cells with ^{125}I and counting the radioactivity on the distal side of the filter or bottom of the dish. *See, e.g., Freshney (1984), supra.*

5

Tumor growth in vivo

Effects of prostate cancer-associated sequences on cell growth can be tested in transgenic or immune-suppressed mice. Knock-out transgenic mice can be made, in which the prostate cancer gene is disrupted or in which a prostate cancer gene is inserted. Knock-
10 out transgenic mice can be made by insertion of a marker gene or other heterologous gene into the endogenous prostate cancer gene site in the mouse genome via homologous recombination. Such mice can also be made by substituting the endogenous prostate cancer gene with a mutated version of the prostate cancer gene, or by mutating the endogenous prostate cancer gene, e.g., by exposure to carcinogens.

15

A DNA construct is introduced into the nuclei of embryonic stem cells. Cells containing the newly engineered genetic lesion are injected into a host mouse embryo, which is re-implanted into a recipient female. Some of these embryos develop into chimeric mice that possess germ cells partially derived from the mutant cell line. Therefore, by breeding the chimeric mice it is possible to obtain a new line of mice containing the introduced genetic
20 lesion (*see, e.g., Capecchi et al., Science 244:1288 (1989)*). Chimeric targeted mice can be derived according to Hogan *et al., Manipulating the Mouse Embryo: A Laboratory Manual*, Cold Spring Harbor Laboratory (1988) and *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, Robertson, ed., IRL Press, Washington, D.C., (1987).

Alternatively, various immune-suppressed or immune-deficient host animals
25 can be used. For example, genetically athymic “nude” mouse (*see, e.g., Giovanella et al., J. Natl. Cancer Inst. 52:921 (1974)*), a SCID mouse, a thymectomized mouse, or an irradiated mouse (*see, e.g., Bradley et al., Br. J. Cancer 38:263 (1978)*; Selby *et al., Br. J. Cancer 41:52 (1980)*) can be used as a host. Transplantable tumor cells (typically about 10^6 cells) injected into isogenic hosts will produce invasive tumors in a high proportions of cases, while
30 normal cells of similar origin will not. In hosts which developed invasive tumors, cells expressing a prostate cancer-associated sequences are injected subcutaneously. After a

suitable length of time, preferably 4-8 weeks, tumor growth is measured (e.g., by volume or by its two largest dimensions) and compared to the control. Tumors that have statistically significant reduction (using, e.g., Student's T test) are said to have inhibited growth.

5 **Methods of identifying variant prostate cancer-associated sequences**

Without being bound by theory, expression of various prostate cancer sequences is correlated with prostate cancer. Accordingly, disorders based on mutant or variant prostate cancer genes may be determined. In one embodiment, the invention provides methods for identifying cells containing variant prostate cancer genes, e.g., determining all or
10 part of the sequence of at least one endogenous prostate cancer genes in a cell. This may be accomplished using any number of sequencing techniques. In a preferred embodiment, the invention provides methods of identifying the prostate cancer genotype of an individual, e.g., determining all or part of the sequence of at least one prostate cancer gene of the individual. This is generally done in at least one tissue of the individual, and may include the evaluation
15 of a number of tissues or different samples of the same tissue. The method may include comparing the sequence of the sequenced prostate cancer gene to a known prostate cancer gene, i.e., a wild-type gene.

The sequence of all or part of the prostate cancer gene can then be compared to the sequence of a known prostate cancer gene to determine if any differences exist. This
20 can be done using any number of known homology programs, such as Bestfit, etc. In a preferred embodiment, the presence of a difference in the sequence between the prostate cancer gene of the patient and the known prostate cancer gene correlates with a disease state or a propensity for a disease state, as outlined herein.

In a preferred embodiment, the prostate cancer genes are used as probes to
25 determine the number of copies of the prostate cancer gene in the genome.

In another preferred embodiment, the prostate cancer genes are used as probes to determine the chromosomal localization of the prostate cancer genes. Information such as chromosomal localization finds use in providing a diagnosis or prognosis in particular when chromosomal abnormalities such as translocations, and the like are identified in the prostate
30 cancer gene locus.

Administration of pharmaceutical and vaccine compositions

In one embodiment, a therapeutically effective dose of a prostate cancer protein or modulator thereof, is administered to a patient. By “therapeutically effective dose” herein is meant a dose that produces effects for which it is administered. The exact dose will depend on the purpose of the treatment, and will be ascertainable by one skilled in the art using known techniques (e.g., Ansel *et al.*, *Pharmaceutical Dosage Forms and Drug Delivery*; Lieberman, *Pharmaceutical Dosage Forms* (vols. 1-3, 1992), Dekker, ISBN 0824770846, 082476918X, 0824712692, 0824716981; Lloyd, *The Art, Science and Technology of Pharmaceutical Compounding* (1999); and Pickar, *Dosage Calculations* (1999)). As is known in the art, adjustments for prostate cancer degradation, systemic versus localized delivery, and rate of new protease synthesis, as well as the age, body weight, general health, sex, diet, time of administration, drug interaction and the severity of the condition may be necessary, and will be ascertainable with routine experimentation by those skilled in the art. U.S. Patent Application N. 09/687,576, further discloses the use of compositions and methods of diagnosis and treatment in prostate cancer is hereby expressly incorporated by reference.

A “patient” for the purposes of the present invention includes both humans and other animals, particularly mammals. Thus the methods are applicable to both human therapy and veterinary applications. In the preferred embodiment the patient is a mammal, preferably a primate, and in the most preferred embodiment the patient is human.

The administration of the prostate cancer proteins and modulators thereof of the present invention can be done in a variety of ways as discussed above, including, but not limited to, orally, subcutaneously, intravenously, intranasally, transdermally, intraperitoneally, intramuscularly, intrapulmonary, vaginally, rectally, or intraocularly. In some instances, e.g., in the treatment of wounds and inflammation, the prostate cancer proteins and modulators may be directly applied as a solution or spray.

The pharmaceutical compositions of the present invention comprise a prostate cancer protein in a form suitable for administration to a patient. In the preferred embodiment, the pharmaceutical compositions are in a water soluble form, such as being present as pharmaceutically acceptable salts, which is meant to include both acid and base addition salts. “Pharmaceutically acceptable acid addition salt” refers to those salts that retain the

biological effectiveness of the free bases and that are not biologically or otherwise undesirable, formed with inorganic acids such as hydrochloric acid, hydrobromic acid, sulfuric acid, nitric acid, phosphoric acid and the like, and organic acids such as acetic acid, propionic acid, glycolic acid, pyruvic acid, oxalic acid, maleic acid, malonic acid, succinic acid, fumaric acid, tartaric acid, citric acid, benzoic acid, cinnamic acid, mandelic acid, methanesulfonic acid, ethanesulfonic acid, p-toluenesulfonic acid, salicylic acid and the like. "Pharmaceutically acceptable base addition salts" include those derived from inorganic bases such as sodium, potassium, lithium, ammonium, calcium, magnesium, iron, zinc, copper, manganese, aluminum salts and the like. Particularly preferred are the ammonium, potassium, sodium, calcium, and magnesium salts. Salts derived from pharmaceutically acceptable organic non-toxic bases include salts of primary, secondary, and tertiary amines, substituted amines including naturally occurring substituted amines, cyclic amines and basic ion exchange resins, such as isopropylamine, trimethylamine, diethylamine, triethylamine, tripropylamine, and ethanolamine.

The pharmaceutical compositions may also include one or more of the following: carrier proteins such as serum albumin; buffers; fillers such as microcrystalline cellulose, lactose, corn and other starches; binding agents; sweeteners and other flavoring agents; coloring agents; and polyethylene glycol.

The pharmaceutical compositions can be administered in a variety of unit dosage forms depending upon the method of administration. For example, unit dosage forms suitable for oral administration include, but are not limited to, powder, tablets, pills, capsules and lozenges. It is recognized that prostate cancer protein modulators (e.g., antibodies, antisense constructs, ribozymes, small organic molecules, *etc.*) when administered orally, should be protected from digestion. This is typically accomplished either by complexing the molecule(s) with a composition to render it resistant to acidic and enzymatic hydrolysis, or by packaging the molecule(s) in an appropriately resistant carrier, such as a liposome or a protection barrier. Means of protecting agents from digestion are well known in the art.

The compositions for administration will commonly comprise a prostate cancer protein modulator dissolved in a pharmaceutically acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers can be used, e.g., buffered saline and the like. These solutions are sterile and generally free of undesirable matter. These compositions may

be sterilized by conventional, well known sterilization techniques. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions such as pH adjusting and buffering agents, toxicity adjusting agents and the like, e.g., sodium acetate, sodium chloride, potassium chloride, calcium chloride, sodium lactate and the like. The concentration of active agent in these formulations can vary widely, and will be selected primarily based on fluid volumes, viscosities, body weight and the like in accordance with the particular mode of administration selected and the patient's needs (e.g., *Remington's Pharmaceutical Science* (15th ed., 1980) and Goodman & Gillman, *The Pharmacological Basis of Therapeutics* (Hardman et al., eds., 1996)).

Thus, a typical pharmaceutical composition for intravenous administration would be about 0.1 to 10 mg per patient per day. Dosages from 0.1 up to about 100 mg per patient per day may be used, particularly when the drug is administered to a secluded site and not into the blood stream, such as into a body cavity or into a lumen of an organ. Substantially higher dosages are possible in topical administration. Actual methods for preparing parenterally administrable compositions will be known or apparent to those skilled in the art, e.g., *Remington's Pharmaceutical Science* and Goodman and Gillman, *The Pharmacological Basis of Therapeutics, supra*.

The compositions containing modulators of prostate cancer proteins can be administered for therapeutic or prophylactic treatments. In therapeutic applications, compositions are administered to a patient suffering from a disease (e.g., a cancer) in an amount sufficient to cure or at least partially arrest the disease and its complications. An amount adequate to accomplish this is defined as a "therapeutically effective dose." Amounts effective for this use will depend upon the severity of the disease and the general state of the patient's health. Single or multiple administrations of the compositions may be administered depending on the dosage and frequency as required and tolerated by the patient. In any event, the composition should provide a sufficient quantity of the agents of this invention to effectively treat the patient. An amount of modulator that is capable of preventing or slowing the development of cancer in a mammal is referred to as a "prophylactically effective dose." The particular dose required for a prophylactic treatment will depend upon the medical condition and history of the mammal, the particular cancer being prevented, as well as other factors such as age, weight, gender, administration route, efficiency, etc. Such prophylactic

treatments may be used, *e.g.*, in a mammal who has previously had cancer to prevent a recurrence of the cancer, or in a mammal who is suspected of having a significant likelihood of developing cancer.

It will be appreciated that the present prostate cancer protein-modulating compounds can be administered alone or in combination with additional prostate cancer modulating compounds or with other therapeutic agent, *e.g.*, other anti-cancer agents or treatments.

In numerous embodiments, one or more nucleic acids, *e.g.*, polynucleotides comprising nucleic acid sequences set forth in Tables 1-16, such as antisense polynucleotides or ribozymes, will be introduced into cells, *in vitro* or *in vivo*. The present invention provides methods, reagents, vectors, and cells useful for expression of prostate cancer-associated polypeptides and nucleic acids using *in vitro* (cell-free), *ex vivo* or *in vivo* (cell or organism-based) recombinant expression systems.

The particular procedure used to introduce the nucleic acids into a host cell for expression of a protein or nucleic acid is application specific. Many procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, spheroplasts, electroporation, liposomes, microinjection, plasma vectors, viral vectors and any of the other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA or other foreign genetic material into a host cell (*see, e.g.*, Berger & Kimmel, *Guide to Molecular Cloning Techniques, Methods in Enzymology* volume 152 (Berger), Ausubel *et al.*, eds., *Current Protocols* (supplemented through 1999), and Sambrook *et al.*, *Molecular Cloning - A Laboratory Manual* (2nd ed., Vol. 1-3, 1989).

In a preferred embodiment, prostate cancer proteins and modulators are administered as therapeutic agents, and can be formulated as outlined above. Similarly, prostate cancer genes (including both the full-length sequence, partial sequences, or regulatory sequences of the prostate cancer coding regions) can be administered in a gene therapy application. These prostate cancer genes can include antisense applications, either as gene therapy (i.e. for incorporation into the genome) or as antisense compositions, as will be appreciated by those in the art.

Prostate cancer polypeptides and polynucleotides can also be administered as vaccine compositions to stimulate HTL, CTL and antibody responses.. Such vaccine

compositions can include, e.g., lipidated peptides (*see, e.g., Vitiello, A. et al., J. Clin. Invest.* 95:341 (1995)), peptide compositions encapsulated in poly(DL-lactide-co-glycolide) ("PLG") microspheres (*see, e.g., Eldridge, et al., Molec. Immunol.* 28:287-294, (1991); Alonso *et al., Vaccine* 12:299-306 (1994); Jones *et al., Vaccine* 13:675-681 (1995)), peptide compositions
5 contained in immune stimulating complexes (ISCOMS) (*see, e.g., Takahashi et al., Nature* 344:873-875 (1990); Hu *et al., Clin Exp Immunol.* 113:235-243 (1998)), multiple antigen peptide systems (MAPs) (*see, e.g., Tam, Proc. Natl. Acad. Sci. U.S.A.* 85:5409-5413 (1988); Tam, *J. Immunol. Methods* 196:17-32 (1996)), peptides formulated as multivalent peptides; peptides for use in ballistic delivery systems, typically crystallized peptides, viral delivery
10 vectors (Perkus, *et al., In: Concepts in vaccine development* (Kaufmann, ed., p. 379, 1996); Chakrabarti, *et al., Nature* 320:535 (1986); Hu *et al., Nature* 320:537 (1986); Kieny, *et al., AIDS Bio/Technology* 4:790 (1986); Top *et al., J. Infect. Dis.* 124:148 (1971); Chanda *et al., Virology* 175:535 (1990)), particles of viral or synthetic origin (*see, e.g., Kofler et al., J. Immunol. Methods.* 192:25 (1996); Eldridge *et al., Sem. Hematol.* 30:16 (1993); Falo *et al., Nature Med.* 7:649 (1995)), adjuvants (Warren *et al., Annu. Rev. Immunol.* 4:369 (1986); Gupta *et al., Vaccine* 11:293 (1993)), liposomes (Reddy *et al., J. Immunol.* 148:1585 (1992); Rock, *Immunol. Today* 17:131 (1996)), or, naked or particle absorbed cDNA (Ulmer, *et al., Science* 259:1745 (1993); Robinson *et al., Vaccine* 11:957 (1993); Shiver *et al., In: Concepts in vaccine development* (Kaufmann, ed., p. 423, 1996); Cease & Berzofsky, *Annu. Rev. Immunol.* 12:923 (1994) and Eldridge *et al., Sem. Hematol.* 30:16 (1993)). Toxin-targeted delivery technologies, also known as receptor mediated targeting, such as those of Avant Immunotherapeutics, Inc. (Needham, Massachusetts) may also be used.

Vaccine compositions often include adjuvants. Many adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide
25 or mineral oil, and a stimulator of immune responses, such as lipid A, *Bordetella pertussis* or *Mycobacterium tuberculosis* derived proteins. Certain adjuvants are commercially available as, e.g., Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or
30 aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides;

polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF, interleukin-2, -7, -12, and other like growth factors, may also be used as adjuvants.

Vaccines can be administered as nucleic acid compositions wherein DNA or
5 RNA encoding one or more of the polypeptides, or a fragment thereof, is administered to a patient. This approach is described, for instance, in Wolff *et al.*, *Science* 247:1465 (1990) as well as U.S. Patent Nos. 5,580,859; 5,589,466; 5,804,566; 5,739,118; 5,736,524; 5,679,647; WO 98/04720; and in more detail below. Examples of DNA-based delivery technologies include “naked DNA”, facilitated (bupivacaine, polymers, peptide-mediated) delivery,
10 cationic lipid complexes, and particle-mediated (“gene gun”) or pressure-mediated delivery (*see, e.g.*, U.S. Patent No. 5,922,687).

For therapeutic or prophylactic immunization purposes, the peptides of the invention can be expressed by viral or bacterial vectors. Examples of expression vectors include attenuated viral hosts, such as vaccinia or fowlpox. This approach involves the use of
15 vaccinia virus, *e.g.*, as a vector to express nucleotide sequences that encode prostate cancer polypeptides or polypeptide fragments. Upon introduction into a host, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits an immune response. Vaccinia vectors and methods useful in immunization protocols are described in, *e.g.*, U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are
20 described in Stover *et al.*, *Nature* 351:456-460 (1991). A wide variety of other vectors useful for therapeutic administration or immunization *e.g.* adeno and adeno-associated virus vectors, retroviral vectors, *Salmonella typhi* vectors, detoxified anthrax toxin vectors, and the like, will be apparent to those skilled in the art from the description herein (*see, e.g.*, Shata *et al.*, *Mol Med Today* 6:66-71 (2000); Shedlock *et al.*, *J Leukoc Biol* 68:793-806 (2000); Hipp *et al.*, *In Vivo* 14:571-85 (2000)).
25

Methods for the use of genes as DNA vaccines are well known, and include placing a prostate cancer gene or portion of a prostate cancer gene under the control of a regulatable promoter or a tissue-specific promoter for expression in a prostate cancer patient. The prostate cancer gene used for DNA vaccines can encode full-length prostate cancer
30 proteins, but more preferably encodes portions of the prostate cancer proteins including peptides derived from the prostate cancer protein. In one embodiment, a patient is

immunized with a DNA vaccine comprising a plurality of nucleotide sequences derived from a prostate cancer gene. For example, prostate cancer-associated genes or sequence encoding subfragments of a prostate cancer protein are introduced into expression vectors and tested for their immunogenicity in the context of Class I MHC and an ability to generate cytotoxic T cell responses. This procedure provides for production of cytotoxic T cell responses against cells which present antigen, including intracellular epitopes.

In a preferred embodiment, the DNA vaccines include a gene encoding an adjuvant molecule with the DNA vaccine. Such adjuvant molecules include cytokines that increase the immunogenic response to the prostate cancer polypeptide encoded by the DNA vaccine. Additional or alternative adjuvants are available.

In another preferred embodiment prostate cancer genes find use in generating animal models of prostate cancer. When the prostate cancer gene identified is repressed or diminished in cancer tissue, gene therapy technology, e.g., wherein antisense RNA directed to the prostate cancer gene will also diminish or repress expression of the gene. Animal models of prostate cancer find use in screening for modulators of a prostate cancer-associated sequence or modulators of prostate cancer. Similarly, transgenic animal technology including gene knockout technology, e.g. as a result of homologous recombination with an appropriate gene targeting vector, will result in the absence or increased expression of the prostate cancer protein. When desired, tissue-specific expression or knockout of the prostate cancer protein may be necessary.

It is also possible that the prostate cancer protein is overexpressed in prostate cancer. As such, transgenic animals can be generated that overexpress the prostate cancer protein. Depending on the desired expression level, promoters of various strengths can be employed to express the transgene. Also, the number of copies of the integrated transgene can be determined and compared for a determination of the expression level of the transgene. Animals generated by such methods find use as animal models of prostate cancer and are additionally useful in screening for modulators to treat prostate cancer.

Kits for Use in Diagnostic and/or Prognostic Applications

For use in diagnostic, research, and therapeutic applications suggested above, kits are also provided by the invention. In the diagnostic and research applications such kits

may include any or all of the following: assay reagents, buffers, prostate cancer-specific nucleic acids or antibodies, hybridization probes and/or primers, antisense polynucleotides, ribozymes, dominant negative prostate cancer polypeptides or polynucleotides, small molecules inhibitors of prostate cancer-associated sequences *etc.* A therapeutic product may include sterile saline or another pharmaceutically acceptable emulsion and suspension base.

In addition, the kits may include instructional materials containing directions (i.e., protocols) for the practice of the methods of this invention. While the instructional materials typically comprise written or printed materials they are not limited to such. Any medium capable of storing such instructions and communicating them to an end user is contemplated by this invention. Such media include, but are not limited to electronic storage media (e.g., magnetic discs, tapes, cartridges, chips), optical media (e.g., CD ROM), and the like. Such media may include addresses to internet sites that provide such instructional materials.

The present invention also provides for kits for screening for modulators of prostate cancer-associated sequences. Such kits can be prepared from readily available materials and reagents. For example, such kits can comprise one or more of the following materials: a prostate cancer-associated polypeptide or polynucleotide, reaction tubes, and instructions for testing prostate cancer-associated activity. Optionally, the kit contains biologically active prostate cancer protein. A wide variety of kits and components can be prepared according to the present invention, depending upon the intended user of the kit and the particular needs of the user. Diagnosis would typically involve evaluation of a plurality of genes or products. The genes will be selected based on correlations with important parameters in disease which may be identified in historical or outcome data.

EXAMPLES

Example 1: Tissue Preparation, Labeling Chips, and Fingerprints

5 Purifying total RNA from tissue sample using TRIzol Reagent

The sample weight is first estimated. The tissue samples are homogenized in 1 ml of TRIzol per 50 mg of tissue using a homogenizer (e.g., Polytron 3100). The size of the generator/probe used depends upon the sample amount. A generator that is too large for the amount of tissue to be homogenized will cause a loss of sample and lower RNA yield. A
10 larger generator (e.g., 20 mm) is suitable for tissue samples weighing more than 0.6 g. Fill tubes should not be overfilled. If the working volume is greater than 2 ml and no greater than 10 ml, a 15 ml polypropylene tube (Falcon 2059) is suitable for homogenization.

Tissues should be kept frozen until homogenized. The TRIzol is added directly to the frozen tissue before homogenization. Following homogenization, the insoluble
15 material is removed from the homogenate by centrifugation at 7500 x g for 15 min. in a Sorvall superspeed or 12,000 x g for 10 min. in an Eppendorf centrifuge at 4°C. The cleared homogenate is then transferred to a new tube(s). Samples may be frozen and stored at -60 to -70°C for at least one month or else continue with the purification.

The next process is phase separation. The homogenized samples are incubated
20 for 5 minutes at room temperature. Then, 0.2 ml of chloroform per 1ml of TRIzol reagent is added to the homogenization mixture. The tubes are securely capped and shaken vigorously by hand (do not vortex) for 15 seconds. The samples are then incubated at room temp. for 2-3 minutes and next centrifuged at 6500 rpm in a Sorvall superspeed for 30 min. at 4°C.

The next process is RNA Precipitation. The aqueous phase is transferred to a
25 fresh tube. The organic phase can be saved if isolation of DNA or protein is desired. Then 0.5 ml of isopropyl alcohol is added per 1ml of TRIzol reagent used in the original homogenization. Then, the tubes are securely capped and inverted to mix. The samples are then incubated at room temp. for 10 minutes and centrifuged at 6500 rpm in Sorvall for 20 min. at 4°C.

The RNA is then washed. The supernatant is poured off and the pellet washed with cold 75% ethanol. 1 ml of 75% ethanol is used per 1 ml of the TRIzol reagent used in the initial homogenization. The tubes are capped securely and inverted several times to loosen pellet without vortexing. They are next centrifuged at <8000 rpm (<7500 x g) for 5 minutes at 4°C.

The RNA wash is decanted. The pellet is carefully transferred to an Eppendorf tube (sliding down the tube into the new tube by use of a pipet tip to help guide it in if necessary). Tube(s) sizes for precipitating the RNA depending on the working volumes. Larger tubes may take too long to dry. Dry pellet. The RNA is then resuspended in an appropriate volume (e.g., 2 -5 ug/ul) of DEPC H₂O. The absorbance is then measured.

The poly A⁺ mRNA may next be purified from total RNA by other methods such as Qiagen's RNeasy kit. The poly A⁺ mRNA is purified from total RNA by adding the oligotex suspension which has been heated to 37°C and mixing prior to adding to RNA. The Elution Buffer is incubated at 70°C. If there is precipitate in the buffer, warm up the 2 x Binding Buffer at 65°C. The the total RNA is mixed with DEPC-treated water, 2 x Binding Buffer, and Oligotex according to Table 2 on page 16 of the Oligotex Handbook and next incubated for 3 minutes at 65°C and 10 minutes at room temperature.

The preparation is centrifuged for 2 minutes at 14,000 to 18,000 g, preferably, at a "soft setting," The supernatant is removed without disturbing Oligotex pellet. A little bit of solution can be left behind to reduce the loss of Oligotex. The supernatant is saved until satisfactory binding and elution of poly A⁺ mRNA has been found.

Then, the preparation is gently resuspended in Wash Buffer OW2 and pipetted onto the spin column and centrifuged at full speed (soft setting if possible) for 1 minute.

Next, the spin column is transferred to a new collection tube and gently resuspended in Wash Buffer OW2 and centrifuged as described herein.

Then, the spin column is transferred to a new tube and eluted with 20 to 100 ul of preheated (70°C) Elution Buffer. The Oligotex resin is gently resuspended by pipetting up and down. The centrifugation is repeated as above and the elution repeated with fresh elution buffer or first eluate to keep the elution volume low.

The absorbance is next read to determine the yield, using diluted Elution Buffer as the blank.

Before proceeding with cDNA synthesis, the mRNA is precipitated before proceeding with cDNA synthesis, as components leftover or in the Elution Buffer from the Oligotex purification procedure will inhibit downstream enzymatic reactions of the mRNA. 0.4 vol. of 7.5 M NH₄OAc + 2.5 vol. of cold 100% ethanol is added and the preparation precipitated at -20°C 1 hour to overnight (or 20-30 min. at -70°C), and centrifuged at 14,000-16,000 x g for 30 minutes at 4°C. Next, the pellet is washed with 0.5 ml of 80% ethanol (-20°C) and then centrifuged at 14,000-16,000 x g for 5 minutes at room temperature. The 80% ethanol wash is then repeated. The last bit of ethanol from the pellet is then dried without use of a speed vacuum and the pellet is then resuspended in DEPC H₂O at 1 µg/ul concentration.

Alternatively the RNA may be purified using other methods (e.g., Qiagen's RNeasy kit).

No more than 100 µg is added to the RNeasy column. The sample volume is adjusted to 100 µl with RNase-free water. 350 µl Buffer RLT and then 250 µl ethanol (100%) are added to the sample. The preparation is then mixed by pipetting and applied to an RNeasy mini spin column for centrifugation (15 sec at >10,000 rpm). If yield is low, reapply the flowthrough to the column and centrifuge again.

Then, transfer column to a new 2 ml collection tube and add 500 µl Buffer RPE and centrifuge for 15 sec at >10,000 rpm. The flowthrough is discarded. 500 µl Buffer RPE and is then added and the preparation is centrifuged for 15 sec at >10,000 rpm. The flowthrough is discarded. and the column membrane dried by centrifuging for 2 min at maximum speed. The column is transferred to a new 1.5-ml collection tube. 30-50 µl of RNase-free water is applied directly onto column membrane. The column is then centrifuged for 1 min at >10,000 rpm and the elution step repeated.

The absorbance is then read to determine yield. If necessary, the material may be ethanol precipitated with ammonium acetate and 2.5X volume 100% ethanol.

First Strand cDNA Synthesis

The first strand can be made using using Gibco's "SuperScript Choice System for cDNA Synthesis" kit. The starting material is 5 ug of total RNA or 1 ug of polyA+ mRNA. For total RNA, 2 ul of SuperScript RT is used; for polyA+ mRNA, 1 ul of SuperScript RT is used. The final volume of first strand synthesis mix is 20 ul. The RNA should be in a volume no greater than 10 ul. The RNA is incubated with 1 ul of 100 pmol T7-T24 oligo for 10 min at 70°C followed by addition on ice of 7 ul of: 4ul 5X 1st Strand Buffer, 2 ul of 0.1M DTT, and 1 ul of 10mM dNTP mix. The preparation is then incubated at 37°C for 2 min before addition of the SuperScript RT followed by incubation at 37°C for 1 hour.

Second Strand Synthesis

For the second strand synthesis, place 1st strand reactions on ice and add: 91 ul DEPC H₂O; 30 ul 5X 2nd Strand Buffer; 3 ul 10mM dNTP mix; 1 ul 10 U/ul E.coli DNA Ligase; 4 ul 10 U/ul E.coli DNA Polymerase; and 1 ul 2 U/ul RNase H. Mix and incubate 2 hours at 16°C. Add 2 ul T4 DNA Polymerase. Incubate 5 min at 16°C. Add 10 ul of 0.5M EDTA.

Cleaning up cDNA

The cDNA is purified using Phenol:Chloroform:Isoamyl Alcohol (25:24:1) and Phase-Lock gel tubes. The PLG tubes are centrifuged for 30 sec at maximum speed. The cDNA mix is then transferred to PLG tube. An equal volume of phenol:chloroform:isamyl alcohol is then added, the preparation shaken vigorously (no vortexing), and centrifuged for 5 minutes at maximum speed. The top aqueous solution is transferred to a new tube and ethanol precipitated by adding 7.5X 5M NH₄OAc and 2.5X volume of 100% ethanol. Next, it is centrifuged immediately at room temperature for 20 min, maximum speed. The supernatant is removed, and the pellet washed with 2X with cold 80% ethanol. As much ethanol wash as possible should be removed before air drying the pellet; and resuspending it in 3 ul RNase-free water.

In vitro Transcription (IVT) and labeling with biotin

In vitro Transcription (IVT) and labeling with biotin is performed as follows:

Pipet 1.5 ul of cDNA into a thin-wall PCR tube. Make NTP labeling mix by combining 2 ul T7 10xATP (75 mM) (Ambion); 2 ul T7 10xGTP (75 mM) (Ambion); 1.5 ul T7 10xCTP (75 mM) (Ambion); 1.5 ul T7 10xUTP (75 mM) (Ambion); 3.75 ul 10 mM Bio-11-UTP (Boehringer-Mannheim/Roche or Enzo); 3.75 ul 10 mM Bio-16-CTP (Enzo); 2 ul 10x T7 transcription buffer (Ambion); and 2 ul 10x T7 enzyme mix (Ambion). The final volume is 20 ul. Incubate 6 hours at 37°C in a PCR machine. The RNA can be further cleaned. Clean-up follows the previous instructions for RNeasy columns or Qiagen's RNeasy protocol handbook. The cRNA often needs to be ethanol precipitated by resuspension in a volume compatible with the fragmentation step.

Fragmentation is performed as follows. 15 ug of labeled RNA is usually fragmented. Try to minimize the fragmentation reaction volume; a 10 ul volume is recommended but 20 ul is all right. Do not go higher than 20 ul because the magnesium in the fragmentation buffer contributes to precipitation in the hybridization buffer. Fragment RNA by incubation at 94 C for 35 minutes in 1 x Fragmentation buffer (5 x Fragmentation buffer is 200 mM Tris-acetate, pH 8.1; 500 mM KOAc; 150 mM MgOAc). The labeled RNA transcript can be analyzed before and after fragmentation. Samples can be heated to 65°C for 15 minutes and electrophoresed on 1% agarose/TBE gels to get an approximate idea of the transcript size range.

For hybridization, 200 ul (10 ug cRNA) of a hybridization mix is put on the chip. If multiple hybridizations are to be done (such as cycling through a 5 chip set), then it is recommended that an initial hybridization mix of 300 ul or more be made. The hybridization mix is: fragment labeled RNA (50 ng/ul final conc.); 50 pM 948-b control oligo; 1.5 pM BioB; 5 pM BioC; 25 pM BioD; 100 pM CRE; 0.1 mg/ml herring sperm DNA; 0.5 mg/ml acetylated BSA; and 300 ul with 1xMES hyb buffer.

The hybridization reaction is conducted with non-biotinylated IVT (purified by RNeasy columns) (see example 1 for steps from tissue to IVT): The following mixture is prepared:

IVT antisense RNA; 4 µg:	µl
Random Hexamers (1 µg/µl):	4 µl
H ₂ O:	<u> µl </u>
	14 µl

- 5 Incubate the above 14 µl mixture at 70°C for 10 min.; then put on ice.

The Reverse transcription procedure uses the following mixture:

0.1 M DTT:	3 µl
50X dNTP mix:	0.6 µl
H ₂ O:	2.4 µl
10 Cy3 or Cy5 dUTP (1mM):	3 µl
SS RT II (BRL):	1 µl
	<u> </u>
	16 µl

The above solution is added to the hybridization reaction and incubated for 30 min., 42°C.

- 15 Then, 1 µl SSII is added and incubated for another hour before being placed on ice.

The 50X dNTP mix contains 25mM of cold dATP, dCTP, and dGTP, 10mM of dTTP and is made by adding 25 µl each of 100mM dATP, dCTP, and dGTP; 10 µl of 100mM dTTP to 15 µl H₂O.]

- 20 RNA degradation is performed as follows. Add 86 µl H₂O, 1.5 µl 1M NaOH/ 2 mM EDTA and incubate at 65°C, 10 min.. For U-Con 30, 500 µl TE/sample spin at 7000 g for 10 min, save flow through for purification. For Qiagen purification, suspend u-con recovered material in 500 µl buffer PB and proceed using Qiagen protocol. For DNase digestion, add 1 ul of 1/100 dilution of DNase/30 ul Rx and incubate at 37°C for 15 min. Incubate at 5 min 95°C to denature the DNase.

25

Sample preparation

For sample preparation, add Cot-1 DNA, 10 µl; 50X dNTPs, 1 µl; 20X SSC, 2.3 µl; Na pyro phosphate, 7.5 µl; 10 mg/ml Herring sperm DNA; 1 ul of 1/10 dilution to 21.8 final vol. Dry in speed vac. Resuspend in 15 µl H₂O. Add 0.38 µl 10% SDS. Heat

95°C, 2 min and slow cool at room temp. for 20 min. Put on slide and hybridize overnight at 64°C. Washing after the hybridization: 3X SSC/0.03% SDS: 2 min., 37.5 ml 20X SSC+0.75ml 10% SDS in 250ml H₂O; 1X SSC: 5 min., 12.5 ml 20X SSC in 250ml H₂O; 0.2X SSC: 5 min., 2.5 ml 20X SSC in 250ml H₂O. Dry slides and scan at appropriate PMT's and channels.

Example 2: Taxol resistant Xenograft Model of Human Prostate Cancer

Treatment regimens that include paclitaxel (Taxol; Bristol-Myers Squibb Company, Princeton, NJ) have been particularly successful in treating hormone-refractory prostate cancer in the phase II setting (Smith et al., Semin. Oncol. 26(1 Suppl 2):109-11 (1999)). However, many patients develop tumors which are initially, or later become, resistant to taxol. To identify genes that may be involved with resistance to taxol, or are regulated in response to taxol resistance, and therefore may be used to treat, or identify, taxol resistance in patients, the following experiments were carried out.

The androgen-independent human cell line CWR22R was grown as a xenograft in nude mice (Nagabhushan et al., Cancer Res. 56(13):3042-3046 (1996); Agus et al., J. Natl. Cancer Inst. 91(21):1869-1876 (1999); Bubendorf et al., J. Natl. Cancer Inst. 91(20):1758-1764 (1999)). Initially, these xenograft tumors were sensitive to therapeutic doses of taxol. The mice were treated continuously with sub-therapeutic doses, and the tumors were allowed to grow for 3-4 weeks, before surgical removal of the tumors. The tumor from an individual mouse was then minced, and a small portion was then injected into a healthy nude mouse, establishing a second

passage of the tumor. This mouse was then treated continuously with the same sub-therapeutic dose of taxol. This process was repeated 14 times, and a portion of each generation of xenograft tumor was collected. There was increasing resistance to therapeutic doses of taxol with each generation. By the end of the process, the tumors were fully resistant to therapeutic doses of taxol. RNA from each generation of tumor was then isolated, and individual mRNA species were quantified using a custom Affymetrix GeneChip® oligonucleotide microarray, with probes to interrogate approximately 35,000

unique mRNA transcripts. Genes were selected that showed a statistically significant up-regulation, or down-regulation, during the subsequent generations of increasingly taxol-resistant tumors. Only one gene was significantly up-regulated, whereas 24 genes were down-regulated; these are presented in Table 10.

The gene sequences identified to be overexpressed in prostate cancer may be used to identify coding regions from the public DNA database. The sequences may be used
5 to either identify genes that encode known proteins, or they may be used to predict the coding regions from genomic DNA using exon prediction algorithms, such as FGENESH (Salamov and Solovyev, 2000, Genome Res. 10:516-522). In addition, one of ordinary skill in the art would understand how to obtain the unigene cluster identification and sequence information according to the exemplar accession numbers provided in Tables 1-16. (see,
10 <http://www.ncbi.nlm.nih.gov/UniGene/>).

15

TABLE1: shows genes, including expression sequence tags, differentially expressed in prostate tumor tissue compared to normal tissue as analyzed using the Affymetrix/Eos Hu01 GeneChip array. Shown are the relative amounts of each gene expressed in prostate tumor samples and various normal tissue samples showing the highest expression of the gene.

10	Pkey:	Unique Eos probeset identifier number		
	ExAccn:	Exemplar Accession number, Genbank accession number		
	UnigeneID:	Unigene number		
	Unigene Title:	Unigene gene title		
	R1:	Ratio of tumor to normal body tissue		
15				
	Pkey	UnigeneID	ExAccn	Unigene Title
				R1
	131919	Hs.272458	AA121266	ESTs
	120328	Hs.290905	AA196979	ESTs; Weakly similar to (define not ava
20	105201	Hs.31412	AA195626	ESTs
	101486	Hs.1852	M24902	acid phosphatase; prostate
	119073	Hs.279477	R32894	ESTs
	133428	Hs.183752	M34376	microseminoprotein; beta-
	128180	Hs.171995	AA595348	kallikrein 3; (prostate specific antigen
25	104080	Hs.57771	AA402971	Homo sapiens mRNA for serine protease (T
	127537	Hs.162859	AA569531	ESTs
	131665	Hs.30343	R22139	ESTs
	101050	Hs.1832	K01911	neuropeptide Y
	130771	Hs.1915	N48056	folate hydrolase (prostate-specific memb
30	108153	Hs.40808	AA054237	ESTs
	107485	Hs.262476	W63793	S-adenosylmethionine decarboxylase 1
	106155	Hs.33287	AA425309	ESTs
	129534	Hs.11260	R73640	ESTs
	100569	Hs.171995	HG2261-HT2351	Antigen, Prostate Specific, Alt. Splice
35	101889	Hs.181350	S39329	kallikrein 2; prostatic
	135389	Hs.99872	U05237	fetal Alzheimer antigen
	101506	Hs.62192	M27436	coagulation factor III (thromboplastin;
	134374	Hs.8236	D62633	ESTs
	133944	Hs.7780	AA045870	ESTs
40	109141	Hs.193380	AA176428	ESTs
	130974	Hs.2178	X57985	H2B histone family; member Q
	114768	Hs.182339	AA149007	ESTs
	104394	Hs.172129	H46617	yp19h1.r1 Soares breast 3NbHBst Homo sap
	125299	Hs.102720	Z39436	ESTs
45	104660	Hs.14846	AA007160	ESTs
	100116	Hs.78045	D00654	actin; gamma 2; smooth muscle; enteric
	131061	Hs.268744	N64328	ESTs; Moderately similar to KIAA0273 [H.
	126645	126645	AI167942	Homo sapiens BAC clone RG041D11 from 7q2
	135153	Hs.95420	N40141	Homo sapiens mRNA for JM27 protein; comp
50	107033	Hs.113314	AA599629	ESTs
	118417		N66048	ESTs; Weakly similar to polymerase [H.sa
	126758	Hs.293960	W37145	ESTs
	115674	Hs.8364	AA406542	ESTs
	134989	Hs.92381	AA236324	ESTs; Weakly similar to !!!! ALU CLASS A
55	107102	Hs.30652	AA609723	ESTs
	116787	Hs.15641	H28581	ESTs
	115719	Hs.59622	AA416997	ESTs
	123209	Hs.203270	AA489711	ESTs
	101664	Hs.121017	M60752	H2A histone family; member A
60	112971	Hs.83883	T17185	ESTs
	102519	Hs.80296	U52969	Purkinje cell protein 4
	117984	Hs.106778	N51919	ESTs
	105840	Hs.22209	AA398533	ESTs
	129523	Hs.274509	M30894	T-cell receptor; gamma cluster
65	132964	Hs.167133	AA031360	ESTs
	121853	Hs.98502	AA425887	ESTs

	115764	Hs.91011	AA421562	anterior gradient 2 (Xenopus laevis; sec	8.9
	119617	Hs.55999	W47380	ESTs	8.9
	100552	Hs.301946	HG2167-HT2237	Protein Kinase Ht31, Camp-Dependent	8.9
5	105627	Hs.23317	AA281245	ESTs	8.8
	101461	Hs.76422	M22430	phospholipase A2; group IIA (platelets;	8.7
	131725	Hs.31146	AA456264	ESTs; Highly similar to (define not ava	8.5
	124526	Hs.293185	N62096	yz61c5.s1 Soares_multiple_sclerosis_2NbH	8.5
	118528	Hs.49397	N67889	ESTs	8.2
	133845	Hs.76704	T68510	ESTs	8.2
10	133354	Hs.334762	AA055552	ESTs; Weakly similar to KIAA0319 [H.sapi	8.1
	105912	Hs.20415	AA402000	ESTs; Weakly similar to GS3786 [H.sapien	8
	119018	Hs.278695	N95796	ESTs	8
	100394	Hs.66052	D84276	CD38 antigen (p45)	8
	114132	Hs.24192	Z38688	ESTs	7.9
15	116786	Hs.301527	H25836	tumor necrosis factor (ligand) superfami	7.7
	106579	Hs.23023	AA456135	ESTs	7.6
	128790	Hs.105700	AA291725	secreted frizzled-related protein 4	7.5
	114965	Hs.72472	AA250737	ESTs	7.4
	112033	Hs.22627	R43162	ESTs	7.1
20	102398		U42359	Human N33 protein form 1 (N33) gene, exo	7
	101201	Hs.2256	L22524	matrix metalloproteinase 7 (matrilysin;	6.9
	109272	Hs.288462	AA195718	ESTs	6.9
	103145	Hs.169849	X66276	myosin-binding protein C; slow-type	6.9
	101803	Hs.155691	M86546	pre-B-cell leukemia transcription factor	6.8
25	120562	Hs.302267	AA280036	ESTs; Weakly similar to W01A6.c [C.elega	6.8
	109112	Hs.257924	AA169379	ESTs	6.8
	109795	Hs.326416	F10707	ESTs	6.7
	107532	Hs.173684	Z19643	ESTs; Weakly similar to (define not ava	6.7
	130336	Hs.171995	X07730	kallikrein 3; (prostate specific antigen	6.6
30	131425	Hs.26691	AA219134	ESTs	6.6
	120588	Hs.16193	AA281591	Homo sapiens mRNA; cDNA DKFZp586B211 (fr	6.6
	132902	Hs.59838	AA490969	ESTs	6.6
	125674	Hs.323378	W28078	H.sapiens mRNA for transmembrane protein	6.6
	133724	Hs.75746	U07919	aldehyde dehydrogenase 6	6.5
35	130343	Hs.278628	AA490262	ESTs; Moderately similar to APXL gene pr	6.5
	120215	Hs.108787	Z41050	Homo sapiens Mcd4p homolog mRNA; complet	6.5
	129215	Hs.126085	AA176867	ESTs	6.5
	131881	Hs.3383	AA010163	upstream regulatory element binding prot	6.5
	133376	Hs.7232	T23670	ESTs	6.4
40	105376	Hs.8768	AA236559	ESTs; Weakly similar to neuronal thread	6.4
	104674	Hs.26289	AA009527	ESTs	6.4
	100727	Hs.334786	X07290	Human HF.12 gene mRNA	6.3
	130150	Hs.15113	AF000573	homogentisate 1,2-dioxygenase (homogenti	6.3
	121770	Hs.278428	AA421714	Homo sapiens mRNA for KIAA0896 protein;	6.3
45	123475	Hs.250528	AA599267	ESTs; Weakly similar to ANKYRIN; BRAIN V	6.3
	133061	Hs.296638	AB000584	prostate differentiation factor	6.3
	116429	Hs.279923	AA609710	ESTs; Weakly similar to similar to GTP-b	6.2
	101233	Hs.878	L29008	sorbitol dehydrogenase	6.2
	104691	Hs.37744	AA011176	ESTs	6.2
50	127248		AA325029	EST27953 Cerebellum II Homo sapiens cDNA	6.2
	127775	Hs.179902	H04106	ESTs; Weakly similar to (define not ava	6.2
	105500	Hs.222399	AA256485	ESTs	6.1
	131463	Hs.2714	X74142	forkhead (Drosophila)-like 1	6.1
	132116	Hs.40289	AA234767	ESTs	6
55	130828	Hs.203213	AA053400	ESTs	5.9
	115357	Hs.72988	AA281793	ESTs	5.8
	105496	Hs.301997	AA256323	ESTs	5.7
	116334	Hs.48948	AA491457	ESTs	5.7
	107968	Hs.61539	AA034020	ESTs	5.7
60	120132	Hs.125019	Z38839	ESTs; Weakly similar to !!!! ALU SUBFAMI	5.6
	106375	Hs.289072	AA443993	ESTs	5.6
	132550	Hs.170195	AA029597	bone morphogenetic protein 7 (osteogenic	5.6
	124777	Hs.140237	R41933	ESTs; Weakly similar to neuronal thread	5.6
	100311	Hs.337616	D50640	phosphodiesterase 3B; cGMP-inhibited	5.6
65	101791	Hs.62354	M83822	Human beige-like protein (BGL) mRNA; par	5.5
	117698	Hs.45107	N41002	ESTs	5.5
	132387	Hs.281434	R70914	heat shock 70kD protein 1	5.5
	122041	Hs.98732	AA431407	Homo sapiens Chromosome 16 BAC clone CIT	5.5
	133723	Hs.262476	AA088851	S-adenosylmethionine decarboxylase 1	5.5

	113938	W81598	ESTs	5.4
	133015	Hs.246315 AA047036	ESTs	5.4
	125745	Hs.75722 Al283493	ribophorin II	5.4
	107295	Hs.80120 T34527	UDP-N-acetyl-alpha-D-galactosamine:polyp	5.4
5	108186	Hs.7780 AA056482	ESTs	5.3
	100184	Hs.21223 D17408	calponin 1; basic; smooth muscle	5.3
	104466	Hs.326392 N25110	Human guanine nucleotide exchange factor	5.3
	104033	Hs.98944 AA365031	ESTs	5.3
10	110844	Hs.167531 N31952	ESTs; Weakly similar to (define not ava	5.3
	129056	Hs.108336 H70627	ESTs; Weakly similar to !!!! ALU SUBFAMI	5.3
	102805	Hs.25351 U90304	iroquois-class homeodomain protein	5.3
	133493	Hs.194369 AA284143	Homo sapiens chromosome 1 atrophin-1 rel	5.3
	129184	Hs.109201 W26769	ESTs; Highly similar to (define not ava	5.2
	134158	Hs.79428 U15174	BCL2/adenovirus E1B 19kD-interacting pro	5.2
15	107240	Hs.159872 D59368	ESTs	5.2
	104787	AA027317	ESTs; Weakly similar to !!!! ALU SUBFAMI	5.2
	123527	Hs.108327 AA608679	damage-specific DNA binding protein 1 (1	5.2
	116646	Hs.194228 F03048	ESTs; Moderately similar to !!!! ALU SUB	5.2
20	101448	Hs.195850 M21389	keratin 5 (epidermolysis bullosa simplex	5.1
	116188	Hs.184598 AA464728	ESTs; Weakly similar to !!!! ALU SUBFAMI	5.1
	126259	Hs.281428 Z21472	ESTs; Moderately similar to !!!! ALU SUB	5.1
	105921	Hs.169119 AA402613	ESTs	5.1
	103375	Hs.54416 X91868	sine oculis homeobox (Drosophila) homolo	5.1
25	128871	Hs.106778 AA400271	ESTs; Highly similar to (define not ava	5.1
	112681	Hs.148932 R87331	ESTs; Moderately similar to semaphorin V	5.1
	105784	Hs.226434 AA350771	ESTs	5.1
	116238	Hs.47144 AA479362	ESTs	5
	102913	Hs.80342 X07696	keratin 15	5
30	103011	Hs.326035 X52541	early growth response 1	5
	126023	H58881	yr36d09.r1 Soares fetal liver spleen 1NF	5
	103709	Hs.13804 AA037316	ESTs	5
	118981	Hs.39288 N93839	ESTs; Weakly similar to !!!! ALU SUBFAMI	5
	134807	Hs.89732 X78932	zinc finger protein 273	5
35	100079	Hs.23311 AB002365	Human mRNA for KIAA0367 gene; partial cd	4.9
	132047	Hs.3796 D83492	EphB6	4.9
	132880	Hs.177537 AA444369	ESTs	4.9
	124049	Hs.74519 F10523	primase; polypeptide 2A (58kD)	4.8
	133330	Hs.71119 U42360	Human N33 mRNA; complete cds	4.8
40	104776	AA026349	ESTs	4.8
	122593	Hs.128749 AA453310	Homo sapiens alpha-methylacyl-CoA racema	4.8
	103912	Hs.143087 AA251078	ESTs	4.8
	113961	Hs.26009 W86307	Homo sapiens mRNA for KIAA0860 protein;	4.8
	105288	Hs.3585 AA233168	ESTs; Weakly similar to coded for by C.	4.8
45	135035	Hs.284186 H89575	ESTs	4.8
	104144	Hs.183390 AA447439	ESTs; Weakly similar to ZINC FINGER PROT	4.8
	129389	Hs.288126 AA621604	ESTs	4.8
	125982	R98091	RAE1 (RNA export 1; S.pombe) homolog	4.8
	125162	Hs.26243 W44682	ESTs	4.8
50	103023	Hs.117950 X53793	multifunctional polypeptide similar to S	4.7
	129735	Hs.32990 W80701	ESTs; Weakly similar to HERV-E envelope	4.7
	104479	Hs.106390 N36040	ESTs	4.7
	103731	AA070545	zm7c3.r1 Stratagene neuroepithelium (#93	4.7
	126575	Hs.127602 W72416	ESTs	4.7
55	124578	Hs.231500 N68321	Human glucose transporter-like protein-I	4.7
	130617	Hs.1674 M90516	glutamine-fructose-6-phosphate transamin	4.7
	116752	Hs.91622 H06373	Homo sapiens clone 24456 mRNA sequence	4.7
	100279	Hs.82007 D42084	Human mRNA for KIAA0094 gene; partial cd	4.7
	126288	Hs.89576 Al479264	ESTs	4.7
	131836	Hs.32990 AA610086	ESTs	4.7
60	106717	Hs.239489 AA465093	TIA1 cytotoxic granule-associated RNA-bi	4.7
	114542	Hs.91011 AA055768	ESTs	4.6
	103806	AA130614	zo1f2.r1 Stratagene neuroepithelium NT2R	4.6
	130529	AA173238	small inducible cytokine A5 (RANTES)	4.6
65	115675	Hs.82065 AA406546	ESTs	4.6
	111386	Hs.293798 N95326	ESTs	4.6
	106503	Hs.29679 AA452411	ESTs	4.6
	119943	Hs.14158 W86835	copine III	4.6
	104459	Hs.100070 M91493	EST	4.6
	100774	Hs.89603 HG371-HT1063	Mucin 1, Epithelial, Alt. Splice 6	4.6

	100652	Hs.142653	HG2825-HT2949	Ret Transforming Gene	4.6
	132015	Hs.3731	D11900	ESTs	4.6
	126086		H70975	yr73g01.r1 Soares fetal liver spleen 1NF	4.6
	130888	Hs.173094	F03819	ESTs	4.6
5	106390	Hs.20166	AA446964	Prostate stem cell antigen	4.6
	126959		AA199853	ESTs; Moderately similar to !!!! ALU SUB	4.5
	131584	Hs.29117	X91648	H.sapiens mRNA for pur alpha extended 3'	4.5
	104838	Hs.20953	AA039481	ESTs	4.5
	125661		R50319	ESTs	4.5
10	103171	Hs.234726	X68733	alpha-1-antichymotrypsin	4.5
	103928	Hs.199160	AA280085	ESTs	4.5
	102899	Hs.75730	X06272	signal recognition particle receptor ('d	4.5
	100892	Hs.180789	HG4557-HT4962	Small Nuclear Ribonucleoprotein U1, 1snr	4.5
	106167	Hs.7956	AA425906	ESTs	4.5
15	129404	Hs.317584	AA172056	ESTs	4.5
	106990	Hs.24758	AA521354	ESTs	4.5
	132316	Hs.44566	U28831	Human protein immuno-reactive with anti-	4.4
	132056	Hs.38176	T89386	Homo sapiens mRNA for KIAA0606 protein;	4.4
	133718	Hs.198760	X15306	neurofilament; heavy polypeptide (200kD)	4.4
20	101470	Hs.1846	M22898	tumor protein p53 (Li-Fraumeni syndrome)	4.4
	131904	Hs.284296	AA143019	ESTs; Highly similar to surface 4 integr	4.4
	105804	Hs.22514	AA383142	ESTs	4.4
	122861	Hs.119394	AA464428	ESTs	4.4
	111336	Hs.29894	N79565	ESTs	4.4
25	121944	Hs.98518	AA429278	ESTs	4.4
	134401	Hs.211577	AA243746	ESTs; Highly similar to CG1 protein [H.s	4.4
	126458	Hs.288969	AA815252	ESTs; Weakly similar to !!!! ALU SUBFAM1	4.4
	133435	Hs.323966	T23983	ESTs; Moderately similar to !!!! ALU SUB	4.4
	105178	Hs.21941	AA187490	ESTs	4.3
30	127315		AA640834	nr27b06.r1 NCI_CGAP_Pr3 Homo sapiens cDN	4.3
	132645	Hs.54424	X87870	H.sapiens mRNA for hepatocyte nuclear fa	4.3
	116162	Hs.282990	AA461487	ESTs; Weakly similar to F52C12.2 [C.eleg	4.3
	118040	Hs.47567	N52876	EST	4.3
35	130008	Hs.278427	M31423	cerebellar degeneration-related protein	4.3
	126607	Hs.114688	W87424	ESTs	4.3
	123061	Hs.105130	AA482030	EST	4.3
	109391	Hs.184245	AA219699	ESTs	4.3
	109175		AA180496	ESTs	4.3
40	127003	Hs.173540	AA550806	ESTs; Weakly similar to (define not ava	4.3
	102547	Hs.46638	U57911	chromosome 11 open reading frame 8	4.3
	134208	Hs.79993	U88871	peroxisomal biogenesis factor 7	4.3
	104258	Hs.5462	AF007216	solute carrier family 4; sodium bicarbon	4.3
	130759	Hs.18946	AA094720	ESTs; Weakly similar to (define not ava	4.3
	132160	Hs.295923	AA281770	seven in absentia (Drosophila) homolog 1	4.3
45	135062	Hs.93872	AA174183	ESTs	4.3
	126510	Hs.334762	R49702	ESTs; Weakly similar to KIAA0319 [H.sapi	4.2
	122055	Hs.98747	AA431732	EST	4.2
	133136	Hs.6574	AF007165	suppressin (nuclear deformed epidermal a	4.2
50	109890	Hs.20843	H04649	ESTs	4.2
	133294	Hs.69997	R79723	H.sapiens mRNA for translin associated z	4.2
	134436	Hs.83190	S80437	fatty acid synthase {3' region} [human,	4.2
	107375	Hs.251064	U88573	NBR2	4.2
	122223	Hs.27413	AA436158	ESTs	4.2
55	103044	Hs.248210	X55777	H.sapiens Mahlavu hepatocellular carcino	4.2
	120125	Hs.59815	W99362	EST	4.2
	128969	Hs.283978	T65327	ESTs; Highly similar to (define not ava	4.2
	129637	Hs.1179	D90359	TATA box binding protein (TBP)-associate	4.2
	106566		AA455921	ESTs; Weakly similar to !!!! ALU SUBFAM1	4.2
	112605	Hs.29852	R79220	ESTs	4.2
60	103364	Hs.279929	X90872	H.sapiens mRNA for gp25L2 protein	4.2
	132811	Hs.57419	U25435	transcriptional repressor	4.2
	126570	Hs.326292	T79274	ESTs	4.2
	116298	Hs.94109	AA489046	ESTs	4.2
65	103024	Hs.105938	X53961	lactotransferrin	4.1
	129133	Hs.108850	R56728	yg95c6.r1 Soares infant brain 1NIB Homo	4.1
	133167	Hs.6641	N98707	kinesin family member 5C	4.1
	126871	Hs.14051	AA351779	ESTs	4.1
	132333	Hs.45032	AA192157	ESTs	4.1
	107376	Hs.327179	U90545	solute carrier family 17 (sodium phospho	4.1

	128517	Hs.100861	AA280617	ESTs; Weakly similar to p60 katanin [H.s	4.1
	130555	Hs.116774	AA450324	ESTs	4.1
	105765	Hs.24183	AA343514	ESTs	4.1
5	126529	Hs.26369	AA133237	ESTs	4.1
	125928	Hs.181889	H29730	ESTs	4.1
	117280	Hs.172129	N22107	ESTs; Moderately similar to !!!! ALU SUB	4.1
	100234	Hs.3085	D29677	KIAA0054 gene product	4.1
	100959	Hs.118127	J00073	actin; alpha; cardiac muscle	4.1
10	107130	Hs.12913	AA620582	ESTs; Weakly similar to (define not ava	4.1
	105035	Hs.8859	AA128486	ESTs	4.1
	126735	Hs.226795	AA808949	glutathione S-transferase pi	4.1
	113056	Hs.8036	T26471	ESTs; Moderately similar to !!!! ALU SUB	4
	102460	Hs.211582	U48959	Homo sapiens myosin light chain kinase (4
15	106968	Hs.26813	AA504631	ESTs; Weakly similar to (define not ava	4
	123107	Hs.104207	AA486071	ESTs	4
	127256	Hs.267967	AA327550	ESTs; Weakly similar to !!!! ALU SUBFAMI	4
	105329	Hs.22862	AA234561	ESTs	4
	115504	Hs.42736	AA291946	ESTs	4
20	120726	Hs.97293	AA293656	ESTs	4
	103576	Hs.94560	Z26317	desmoglein 2	4
	127889	Hs.144941	AI147408	ESTs	4
	106394	Hs.25320	AA447223	ESTs	4
	128046		AA873285	ESTs	4
25	103391	Hs.114366	X94453	pyrroline-5-carboxylate synthetase (glut	4
	106448	Hs.27004	AA449455	ESTs	4
	126513	Hs.86276	W27601	ESTs; Moderately similar to (define not	4
	129593	Hs.98314	AA487015	ESTs; Weakly similar to !!!! ALU SUBFAMI	3.9
	110151	Hs.31608	H18836	ESTs	3.9
30	105344	Hs.8645	AA235303	ESTs	3.9
	104791	Hs.301871	AA029046	ESTs	3.9
	123442	Hs.111496	AA598803	ESTs	3.9
	127800	Hs.79428	AA521047	BCL2/adenovirus E1B 19kD-interacting pro	3.9
	114555	Hs.167904	AA058594	ESTs	3.9
35	122138	Hs.163960	AA435549	ESTs	3.9
	129565	Hs.198726	X77777	vasoactive intestinal peptide receptor 1	3.9
	103471	Hs.75216	Y00815	protein tyrosine phosphatase; receptor t	3.9
	133908	Hs.325474	M83216	caldesmon 1	3.9
	105635	Hs.301985	AA281508	ESTs	3.9
40	134285	Hs.81086	AA460012	solute carrier family 22 (organic cation	3.9
	134125	Hs.50421	R38102	KIAA0203 gene product	3.9
	125628	Hs.241493	AA418069	natural killer-tumor recognition sequenc	3.9
	103695	Hs.186600	AA018758	ESTs	3.9
	100642	Hs.182183	HG2743-HT3926	Caldesmon 1, Alt. Splice 6, Non-Muscle	3.9
45	104334	Hs.78771	D82614	ESTs	3.9
	110242	Hs.19978	H26417	ESTs	3.9
	125298	Hs.289008	Z39255	ESTs	3.9
	104060	Hs.303193	AA397968	zt87a9.r1 Soares_testis_NHT Homo sapiens	3.9
	105823	Hs.293960	AA398197	ESTs	3.9
50	126499	Hs.110445	AA315671	ESTs; Moderately similar to unknown [M.m	3.9
	130752	Hs.18895	D50927	KIAA0137 gene product	3.8
	123494	Hs.112110	AA599786	ESTs	3.8
	104846	Hs.32478	AA040154	ESTs	3.8
	108921	Hs.71721	AA142913	ESTs	3.8
55	115506	Hs.45207	AA292537	ESTs	3.8
	100452	Hs.241552	D87742	Human mRNA for KIAA0268 gene; partial cd	3.8
	104454	Hs.129228	M84443	galactokinase 2	3.8
	108730	Hs.102859	AA126254	ESTs	3.8
	131223	Hs.24427	AA247788	ESTs; Highly similar to (define not ava	3.8
60	104784	Hs.269228	AA027055	ESTs	3.8
	104946	Hs.73848	AA069549	ESTs	3.8
	106932	Hs.9394	AA495926	ESTs	3.8
	101724	Hs.620	M69225	bullous pemphigoid antigen 1 (230/240kD)	3.8
	106140	Hs.14912	AA424524	Homo sapiens mRNA for KIAA0286 gene; par	3.8
65	128135	Hs.269721	AA913491	ESTs	3.8
	120030	Hs.58694	W92051	ESTs	3.8
	126457	Hs.50382	AA007489	zh98g04.r1 Soares_fetal_liver_spleen_1NF	3.8
	123917	Hs.112969	AA621311	EST	3.7
	110714	Hs.17752	H95978	Homo sapiens phosphatidylserine-specific	3.7
	130577	Hs.162	M35410	insulin-like growth factor binding prote	3.7

	117667	Hs.44708	N39214	ser-Thr protein kinase related to the my	3.7
	126104	Hs.39712	N77278	ESTs; Weakly similar to BONE/CARTILAGE P	3.7
	100379	Hs.278721	D82060	Homo sapiens mRNA for membrane protein w	3.7
5	115646	Hs.305971	AA404352	ESTs	3.7
	125792	Hs.193700	AI005388	ESTs; Moderately similar to !!!! ALU SUB	3.7
	102162	Hs.1592	U18291	CDC16 (cell division cycle 16; S. cerevi	3.7
	128530	Hs.183475	AA504343	ESTs; Moderately similar to !!!! ALU SUB	3.7
	119940	Hs.272531	W86779	EST	3.7
10	110769	Hs.23837	N22222	yw34b06.s1 Morton Fetal Cochlea Homo sap	3.7
	132914	Hs.60293	AA496037	ESTs	3.7
	113594	Hs.15683	T92030	ESTs	3.7
	103702	Hs.279952	AA027793	ESTs; Highly similar to (define not ava	3.7
	130780	Hs.19347	AA248406	ESTs	3.7
	123288	Hs.291025	AA495836	EST	3.7
15	120691	Hs.22380	AA291173	ESTs	3.7
	103153	Hs.75295	X66534	guanylate cyclase 1; soluble; alpha 3	3.7
	129201	Hs.109390	H19969	ESTs	3.7
	114798	Hs.54900	AA159181	ESTs	3.7
	126801	Hs.7337	AA512902	ESTs	3.7
20	105503	Hs.31707	AA256616	ESTs	3.7
	104260	Hs.194283	AF008192	Homo sapiens putative GR6 protein (GR6)	3.7
	125980	Hs.35699	R97219	ESTs	3.7
	123255	Hs.105273	AA490890	ESTs	3.6
	103862	Hs.6363	AA206625	ESTs	3.6
25	100696	Hs.121686	HG3162-HT3339	Transcription Factor lia	3.6
	134917	Hs.166994	X87241	FAT tumor suppressor (Drosophila) homolo	3.6
	103520		Y10511	H.sapiens mRNA for CD176 protein	3.6
	113778	Hs.302738	W15263	ESTs	3.6
	101838	Hs.75511	M92934	connective tissue growth factor	3.6
30	113702		T97307	ESTs; Moderately similar to !!!! ALU SUB	3.6
	118201	Hs.48428	N59800	EST	3.6
	116519	Hs.68554	C20780	EST	3.6
	105886	Hs.22983	AA400517	ESTs; Moderately similar to UDP-GLUCOSE:	3.6
	106709	Hs.170291	AA464696	ESTs	3.6
35	127858	Hs.27973	AA806365	oc26h07.s1 NCL_CGAP_GCB1 Homo sapiens cD	3.6
	101964		S81578	dioxin-responsive gene {putative polyade	3.6
	105508	Hs.326416	AA256680	ESTs	3.6
	116844	Hs.337434	H64938	ESTs	3.6
	105372	Hs.142296	AA236481	ESTs	3.6
40	100745	Hs.144630	HG3510-HT3704	V-Erba Related Ear-3 Protein	3.6
	127521	Hs.164018	AA809982	ESTs	3.6
	110758	Hs.274265	N21365	talin	3.6
	107307	Hs.44155	T52099	creatine kinase; mitochondrial 2 (sarcom	3.6
	133200	Hs.183639	AA432248	ESTs	3.6
45	114774	Hs.184325	AA150043	ESTs	3.6
	120265	Hs.270696	AA173759	ESTs; Moderately similar to !!!! ALU SUB	3.6
	134359	Hs.199067	M34309	v-erb-b2 avian erythroblastic leukemia v	3.6
	116250	Hs.44829	AA480975	ESTs; Moderately similar to !!!! ALU SUB	3.6
	106313	Hs.35841	AA436459	nuclear factor I/X (CCAAT-binding transc	3.6
50	131898	Hs.279780	N52232	ESTs	3.6
	133444	Hs.73793	M27281	vascular endothelial growth factor	3.6
	128232	Hs.334641	H06296	ESTs	3.6
	135357	Hs.79572	AA235803	ESTs	3.5
	457951		AI369384	arylsulfatase D	3.5
55	108407		AA075519	zm87h9.s1 Stratagene ovarian cancer (#93	3.5
	126659		T16245	a disintegrin and metalloproteinase doma	3.5
	104189	Hs.301804	AA485805	ESTs	3.5
	125956	Hs.129014	N53276	ESTs	3.5
	103026	Hs.79386	X54162	Human mRNA for a 64 Kd autoantigen expre	3.5
60	133011	Hs.171921	AA042990	sema domain; immunoglobulin domain (Ig);	3.5
	131379	Hs.26176	R49035	ESTs	3.5
	126742	Hs.169359	H64106	yr57e06.r1 Soares fetal liver spleen 1NF	3.5
	105560	Hs.306915	AA262783	ESTs	3.5
	118472	Hs.42179	N68818	ESTs	3.5
65	105623	Hs.30127	AA280895	ESTs; Highly similar to !!!! ALU SUBFAMI	3.5
	120262	Hs.145807	AA172076	ESTs; Moderately similar to !!!! ALU SUB	3.5
	105027	Hs.26771	AA126472	ESTs	3.5
	130760	Hs.18953	AA128997	phosphodiesterase 9A	3.5
	117473	Hs.155560	N30157	ESTs	3.5

	102663	Hs.168075	U70322	karyopherin (importin) beta 2	3.5
	126349	Hs.13531	AA442868	ESTs; Weakly similar to (define not ava	3.5
	132154	Hs.41119	N67179	ESTs	3.5
5	131689	Hs.30696	AA599653	transcription factor-like 5 (basic helix	3.5
	127862	Hs.163191	AA765305	EST	3.5
	126995	Hs.189810	W26950	Human DNA sequence from PAC 388M5 on chr	3.5
	119071		R31180	ESTs	3.5
	103941	Hs.96593	AA282978	ESTs	3.5
	110721	Hs.31319	H97678	ESTs	3.5
10	126586	Hs.43086	AA011247	ESTs	3.5
	103106	Hs.1857	X62025	phosphodiesterase 6G; cGMP-specific; rod	3.5
	116357	Hs.90797	AA504806	Homo sapiens clone 23620 mRNA sequence	3.5
	105309	Hs.4104	AA233790	ESTs	3.5
15	130796	Hs.19525	R39390	ESTs	3.5
	109101	Hs.52184	AA167708	ESTs	3.5
	103134	Hs.2839	X65724	Norrie disease (pseudoglioma)	3.5
	131798	Hs.301449	X86098	adenovirus 5 E1A binding protein	3.5
	118535	Hs.49418	N67968	ESTs	3.5
	102592	Hs.11223	U62389	Human putative cytosolic NADP-dependent	3.4
20	125905	Hs.6456	T69868	chaperonin containing TCP1; subunit 2 (b	3.4
	109160	Hs.301997	AA179387	ESTs	3.4
	105327	Hs.211593	AA234440	ESTs	3.4
	106586	Hs.57787	AA456598	ESTs	3.4
	122635		AA454085	EST	3.4
25	132413	Hs.260116	AA132969	metalloprotease 1 (pitriylsin family)	3.4
	131938	Hs.34956	AA283620	ESTs	3.4
	133871	Hs.182793	AA454597	ESTs	3.4
	107175	Hs.292503	AA621751	ESTs; Weakly similar to KIAA0601 protein	3.4
	101188	Hs.184298	L20320	cyclin-dependent kinase 7 (homolog of Xe	3.4
30	126422	Hs.237658	H48518	ESTs; Highly similar to apolipoprotein A	3.4
	118475		N66845	ESTs; Weakly similar to !!!! ALU CLASS B	3.4
	104558	Hs.88959	R56678	ESTs; Weakly similar to !!!! ALU SUBFAM1	3.4
	128307	Hs.132005	AI453794	ESTs	3.4
	112254	Hs.25829	R51831	ESTs	3.4
35	125408	Hs.89578	N72353	yv37e12.r1 Soares fetal liver spleen 1NF	3.4
	109834	Hs.175955	H00604	ESTs	3.4
	130844	Hs.20191	D12122	seven in absentia (Drosophila) homolog 2	3.4
	127143	Hs.20843	AA533553	nj68h04.s1 NCL_CGAP_Pr10 Homo sapiens cD	3.4
	135309	Hs.42500	D25984	ESTs	3.4
40	125724	Hs.295978	AA083407	stimulated trans-acting factor (50 kDa)	3.4
	127692	Hs.187983	AI021912	ESTs	3.4
	116674	Hs.92127	F04816	ESTs	3.4
	134700	Hs.8868	AA481414	golgi SNAP receptor complex member 1	3.4
	114846	Hs.166196	AA234929	ESTs	3.4
45	103649	Hs.155983	Z70219	H.sapiens mRNA for 5'UTR for unknown pro	3.4
	134835	Hs.89925	L04569	calcium channel; voltage-dependent; L ty	3.4
	130568	Hs.16085	AA232535	ESTs; Highly similar to (define not ava	3.4
	111331	Hs.15978	N78773	ESTs	3.4
	106036	Hs.10653	AA412505	ESTs	3.4
50	130987	Hs.21893	R45698	ESTs	3.4
	112814	Hs.35828	R98192	ESTs	3.4
	127815	Hs.255015	AA876009	ob93c10.s1 NCL_CGAP_GCB1 Homo sapiens cD	3.4
	100144	Hs.75616	D13643	KIAA0018 gene product	3.4
	101129	Hs.247992	L10405	Homo sapiens DNA binding protein for sur	3.4
55	130874	Hs.20621	T08287	ESTs	3.4
	106882	Hs.26994	AA489009	ESTs	3.4
	103855	Hs.302267	AA195179	ESTs	3.4
	125957		H45213	yo03b08.r1 Soares adult brain N2b5HB55Y	3.3
	114048	Hs.146085	W94613	ESTs	3.3
60	109826	Hs.75354	F13702	ESTs	3.3
	125355	Hs.170098	R45630	ESTs; Highly similar to KIAA0372 [H.sapi	3.3
	104182	Hs.143792	AA479990	ESTs; Weakly similar to glioma amplified	3.3
	100294	Hs.75454	D49396	Human mRNA for Apo1_Human (MER5(Aop1-Mou	3.3
	131688	Hs.30692	U24153	p21 (CDKN1A)-activated kinase 2	3.3
65	116256	Hs.88201	AA481256	ESTs; Weakly similar to (define not ava	3.3
	102034	Hs.230	U05291	fibromodulin	3.3
	130072	Hs.14658	R99606	Human chromosome 5q13.1 clone 5G8 mRNA	3.3
	114615	Hs.159456	AA083812	ESTs; Highly similar to (define not ava	3.3
	128707	Hs.104105	AA136474	Meis (mouse) homolog 2	3.3

	115048	Hs.190057	AA252668	ESTs	3.3
	125862	Hs.31110	H12084	ESTs	3.3
	135142	Hs.24192	R31679	ESTs	3.3
5	103119	Hs.2877	X63629	cadherin 3; P-cadherin (placental)	3.3
	104460	Hs.62604	M91504	ESTs	3.3
	100365	Hs.79284	D78611	mesoderm specific transcript (mouse) hom	3.3
	131524	Hs.301804	N39152	ESTs	3.3
	102165	Hs.159627	U18321	Death associated protein 3	3.3
10	126966	Hs.182575	R38438	solute carrier family 15 (H+/peptide tra	3.3
	124839	Hs.140942	R55784	ESTs	3.3
	100709	Hs.100469	HG3264-HT3441	Af-6 (Gb:U02478)	3.3
	132967	Hs.61635	AA032221	Homo sapiens BAC clone RG041D11 from 7q2	3.3
	102927	Hs.65114	X12876	keratin 18	3.3
15	132616	Hs.283558	AA386264	ESTs	3.3
	125132	Hs.129781	W15495	ESTs	3.3
	111225	Hs.31652	N68989	ESTs	3.3
	114956	Hs.87113	AA243681	ESTs	3.3
	122235	Hs.112227	AA436475	ESTs	3.3
	112325	Hs.12315	R56055	ESTs	3.3
20	123360	Hs.178604	AA504784	ESTs	3.3
	105150	Hs.155995	AA169640	Homo sapiens mRNA for KIAA0643 protein;	3.3
	107391	Hs.284294	W02877	ESTs	3.3
	113058	Hs.7569	T26893	EST	3.3
25	134371	Hs.82318	S69790	Brush-1	3.3
	125669	Hs.333256	R51308	ESTs; Moderately similar to !!!! ALU SUB	3.3
	111506	Hs.294105	R07726	ESTs	3.3
	122974	Hs.194215	AA478625	ESTs	3.3
	102369	Hs.299867	U39840	hepatocyte nuclear factor 3; alpha	3.3
	120408	Hs.190151	AA235045	ESTs	3.3
30	117993	Hs.47402	N52039	ESTs; Weakly similar to !!!! ALU SUBFAMI	3.3
	129586	Hs.11500	AA437118	ESTs	3.3
	128138	Hs.126494	AI200825	ESTs	3.3
	127265		AA332751	EST37214 Embryo, 8 week I Homo sapiens c	3.3
35	107674	Hs.41143	AA011027	Homo sapiens mRNA for KIAA0581 protein;	3.2
	104866	Hs.293691	AA045342	ESTs	3.2
	103427	Hs.250655	X97303	H.sapiens mRNA for Ptg-12 protein	3.2
	132990	Hs.334334	AA458761	ESTs	3.2
	127017	Hs.251946	AA740146	ESTs	3.2
40	132313	Hs.44481	U13220	forkhead (Drosophila)-like 6	3.2
	106880	Hs.32425	AA488889	ESTs	3.2
	107039	Hs.169780	AA599751	homologous to yeast nitrogen permease (c	3.2
	120870	Hs.292581	AA357172	ESTs	3.2
	107920	Hs.284207	AA027951	ESTs	3.2
45	104165	Hs.105116	AA459160	EST	3.2
	107012	Hs.63908	AA598745	ESTs	3.2
	103605	Hs.194657	Z35402	H.sapiens gene encoding E-cadherin, exon	3.2
	124006	Hs.270016	D60302	ESTs	3.2
	101300	Hs.74137	L40391	Homo sapiens (clone s153) mRNA fragment	3.2
	101183	Hs.795	L19779	H2A histone family; member O	3.2
50	125596		R25698	yg44h11.r2 Soares infant brain 1NIB Homo	3.2
	127261		AA661567	nu86b02.s1 NCI_CGAP_Alvi Homo sapiens cD	3.2
	120090	Hs.59554	W94591	ESTs	3.2
	129393	Hs.166982	D13435	phosphatidylinositol glycan; class F	3.2
55	120923	Hs.97129	AA382283	ESTs	3.2
	118907	Hs.274256	N91003	ESTs	3.2
	111552	Hs.191185	R09411	ESTs	3.2
	104431	Hs.99913	J03019	adrenergic; beta-1-; receptor	3.2
	133551	Hs.278634	D63480	Human mRNA for KIAA0146 gene; partial cd	3.2
60	131615	Hs.192803	D14533	xeroderma pigmentosum; complementation g	3.2
	126547	Hs.84072	U47732	transmembrane 4 superfamily member 3	3.2
	103172	Hs.116774	X68742	integrin; alpha 1	3.2
	113867	Hs.24095	W68845	ESTs	3.2
	133323	Hs.70937	Z83735	H3 histone family; member K	3.2
	111597	Hs.189716	R11499	ESTs	3.2
65	121515	Hs.104696	AA412133	ESTs	3.2
	107445	Hs.6639	W28406	ESTs	3.2
	106887	Hs.334335	AA489091	ESTs	3.2
	123052	Hs.185766	AA481806	ESTs	3.2
	107072	Hs.130760	AA609113	Homo sapiens mRNA; cDNA DKFZp586N0318 (f	3.2

	102214	Hs.32964	U23752	SRY (sex-determining region Y)-box 11	3.2
	123147		AA487961	ab11h6.s1 Stratagene lung (#93721) Homo	3.2
	125435	Hs.272138	R00940	ye87g03.r1 Soares fetal liver spleen 1NF	3.2
5	116246	Hs.250646	AA479961	ESTs; Highly similar to ubiquitin-conjug	3.2
	105169	Hs.180789	AA180321	Homo sapiens (clone S164) mRNA; 3' end o	3.2
	134001	Hs.78344	AF001548	myosin; heavy polypeptide 11; smooth mus	3.2
	124866	Hs.304389	R68571	ESTs	3.2
	133205	Hs.67619	AA089559	Homo sapiens mRNA; chromosome 1 specific	3.2
10	102986	Hs.182378	X17648	colony stimulating factor 1 (macrophage)	3.2
	101232	Hs.242894	L28997	ADP-ribosylation factor-like 1	3.1
	132906	Hs.234896	AA142857	ESTs; Highly similar to geminin [H.sapie	3.1
	104281	Hs.5669	C14290	ESTs	3.1
	123926	Hs.227933	AA621348	ESTs; Highly similar to (define not ava	3.1
	134464	Hs.239720	N79354	ESTs; Weakly similar to Rga [D.melanogas	3.1
15	105322	Hs.16346	AA234100	ESTs	3.1
	100631	Hs.48332	HG2709-HT2805	Serine/Threonine Kinase (Gb:Z25431)	3.1
	130791	Hs.199263	AA259102	ESTs; Highly similar to (define not ava	3.1
	131220	Hs.300855	R77200	ESTs	3.1
	113237	Hs.123642	T62857	ESTs	3.1
20	125562	Hs.98968	AI494372	ESTs	3.1
	134110	Hs.79136	U41060	Human breast cancer; estrogen regulated	3.1
	132393	Hs.47334	W85888	ESTs; Moderately similar to !!!! ALU SUB	3.1
	107439	Hs.296842	W27995	ESTs; Moderately similar to non-muscle m	3.1
25	125863	Hs.40719	AA299096	Homo sapiens mRNA; cDNA DKFZp564M0916 (f	3.1
	105811	Hs.286192	AA394121	ESTs	3.1
	129284	Hs.296141	AA104023	ESTs	3.1
	125321	Hs.178294	T86652	ESTs	3.1
	107332	Hs.183297	T87750	ESTs	3.1
30	123570	Hs.109653	AA608955	ESTs	3.1
	100384	Hs.90800	D83646	matrix metalloproteinase 16 (membrane-in	3.1
	109063	Hs.38972	AA161043	tetraspan 1	3.1
	133284	Hs.182828	U09367	zinc finger protein 136 (clone pHZ-20)	3.1
	131839	Hs.33010	H80622	Homo sapiens mRNA for KIAA0633 protein;	3.1
35	117606	Hs.44698	N35115	ESTs	3.1
	418998	Hs.287849	F13215	ESTs	3.1
	125180	Hs.103120	W58344	ESTs	3.1
	100789		HG3893-HT4163	Phosphoglucomutase 1, Alt. Splice	3.1
	126017	Hs.159440	H60487	ESTs	3.1
40	132452	Hs.247324	AA005262	Homo sapiens DNA sequence from PAC 262D1	3.1
	129077	Hs.108479	H78886	ESTs	3.1
	126563	Hs.181368	W26247	U5 snRNP-specific protein (220 kD); orth	3.1
	129650	Hs.118258	N52554	ESTs	3.1
	123465		AA599033	ESTs	3.1
45	126486	Hs.152316	AA345339	EST51345 Gall bladder II Homo sapiens cD	3.1
	126460	Hs.167031	W01616	za36d05.r1 Soares fetal liver spleen 1NF	3.1
	118697	Hs.43234	N72094	ESTs	3.1
	103860	Hs.38057	AA203742	ESTs	3.1
	127968	Hs.124347	AA971439	ESTs	3.1
50	124984	Hs.223241	T47566	yb15c11.s1 Stratagene placenta (#937225)	3.1
	103903	Hs.15220	AA249334	j312.seq.F Human fetal heart, Lambda ZAP	3.1
	106697	Hs.22242	AA463737	ESTs	3.1
	130892	Hs.20993	AA442604	ESTs; Weakly similar to Ydr374cp [S.cere	3
	114032	Hs.35014	W92779	ESTs	3
55	128835	Hs.106390	W15528	ESTs	3
	103667	Hs.247815	Z80788	H.sapiens H4/I gene	3
	126264	Hs.250614	N42897	yy13h06.r1 Soares melanocyte 2NbHM Homo	3
	132626	Hs.21275	D25755	ESTs	3
	131107	Hs.75354	N87590	ESTs	3
	126780	Hs.5811	R12421	ESTs	3
60	127363	Hs.22116	AA307744	Homo sapiens Cdc14B1 phosphatase mRNA; c	3
	103690	Hs.84063	AA016186	ESTs	3
	102589	Hs.8867	U62015	Homo sapiens Cyr61 mRNA, complete cds	3
	125144	Hs.24336	W37999	ESTs	3
65	132977	Hs.301404	U28666	RNA binding motif protein 3	3
	120714	Hs.146170	AA292689	ESTs	3
	101038	Hs.79411	J05249	replication protein A2 (32kD)	3
	102856	Hs.248177	X00090	Human histone H3 gene	3
	105516	Hs.30738	AA257971	ESTs	3
	131137	Hs.33287	U85193	nuclear factor I/B	3

	127221	Hs.241551	AI354332	ESTs	3
	411888	Hs.24104	R26708	ESTs	3
	131684	Hs.3066	U26174	granzyme K (serine protease; granzyme 3;	3
5	100629	Hs.21291	HG2706-HT2802	Serine/Threonine Kinase (Gb:Z25428)	3
	119944	Hs.58915	W86838	EST	3
	113801	Hs.118281	W38418	zinc finger protein 266	3
	133780	Hs.76152	M14219	decorin	3
	104690	Hs.14449	AA010889	ESTs	3
	126371	Hs.304139	N57645	EST	3
10	127635	Hs.116346	AA766903	ESTs	3
	128434	Hs.143880	AI190914	ESTs	3
	435761	Hs.187555	AA701941	ESTs	3
	125025	Hs.50748	T71561	ESTs	3
	124940	Hs.103804	R99599	heterogeneous nuclear ribonucleoprotein	3
15	128742	Hs.251531	D00783	proteasome (prosome; macropain) subunit;	3
	107147	Hs.10450	AA621125	Homo sapiens chromosome 2; 10 repeat reg	3
	112068	Hs.22545	R43910	ESTs	3
	105346	Hs.263727	AA235465	ESTs; Moderately similar to !!!! ALU SUB	3
	130972	Hs.21739	AA370302	Homo sapiens mRNA; cDNA DKFZp58611518 (f	3
20	131230	Hs.274407	AA149987	thymus specific serine peptidase	3
	133743	Hs.75847	N79435	ESTs	3
	127402	Hs.227949	AA358869	ESTs; Highly similar to SEC13-RELATED PR	3
	117483	Hs.44189	N30426	ESTs	3
	123659	Hs.112699	AA609368	ESTs	3
25	103963	Hs.63290	AA298588	EST114219 HSC172 cells II Homo sapiens c	3
	103795	Hs.7367	AA112222	ESTs; Moderately similar to (define not	3
	115092	Hs.80975	AA255903	CD39-like 4	2.9
	134831	Hs.89890	S72370	pyruvate carboxylase	2.9
	128579	Hs.101810	AA093378	ESTs; Weakly similar to !!!! ALU SUBFAMI	2.9
30	134193	Hs.7980	F09570	ESTs	2.9
	123522	Hs.112575	AA608577	ESTs	2.9
	107109	Hs.32793	AA609943	ESTs	2.9
	134694	Hs.88556	D50405	histone deacetylase 1	2.9
	134399	Hs.82689	H99801	tumor rejection antigen (gp96) 1	2.9
35	134632	Hs.174139	AA398710	H. sapiens RNA for CLCN3	2.9
	106683	Hs.14512	AA461495	ESTs	2.9
	108555		AA084963	zn13e12.s1 Stratagene hNT neuron (#93723	2.9
	100953	Hs.2110	HG945-HT945	Nucleic Acid-Binding Protein (Gb:L12693)	2.9
	130597	Hs.16492	AA173998	ESTs; Weakly similar to weakly similar t	2.9
40	101813	Hs.139226	M87338	replication factor C (activator 1) 2 (40	2.9
	106636	Hs.286	AA459950	ESTs	2.9
	129109	Hs.108708	AA491295	calcium/calmodulin-dependent protein kin	2.9
	125819	Hs.251871	AA044840	stromal cell-derived factor 1	2.9
	106282	Hs.9857	AA433946	ESTs; Weakly similar to (define not ava	2.9
45	100386	Hs.301636	D83703	peroxisomal biogenesis factor 6	2.9
	114546	Hs.98074	AA056263	ESTs; Moderately similar to !!!! ALU SUB	2.9
	105914	Hs.9701	AA402224	Homo sapiens growth arrest and DNA-damag	2.9
	108552		AA084912	zn11c7.s1 Stratagene hNT neuron (#937233	2.9
	126505	Hs.190057	W26894	16a11 Human retina cDNA randomly primed	2.9
50	134098	Hs.79086	X06323	Human MRL3 mRNA for ribosomal protein L3	2.9
	129721	Hs.211539	L19161	eukaryotic translation initiation factor	2.9
	100076	Hs.277422	AB000897	Homo sapiens mRNA for cadherin FIB3, par	2.9
	117466	Hs.44104	N29862	ESTs	2.9
	106335	Hs.36688	AA437258	ESTs; Moderately similar to WAP four-dis	2.9
55	134510	Hs.250870	U25265	protein kinase; mitogen-activated; kinas	2.9
	105835	Hs.32995	AA398412	ESTs	2.9
	106611	Hs.26267	AA458904	ESTs; Weakly similar to torsinA [H.sapie	2.9
	134087	Hs.173824	U51166	thymine-DNA glycosylase	2.9
	100641	Hs.182183	HG2743-HT2846	Caldesmon 1, Alt. Splice 4, Non-Muscle	2.9
60	104602		R86920	ESTs	2.9
	117203	Hs.42738	H99799	ESTs	2.9
	131889	Hs.34073	AA401912	BH-protocadherin (brain-heart)	2.9
	101707	Hs.155212	M65131	methylmalonyl Coenzyme A mutase	2.9
	115271	Hs.5724	AA279422	ESTs	2.9
65	125812	Hs.287912	H73420	lectin; mannose-binding; 1	2.9
	110740	Hs.19762	H99675	ESTs	2.9
	103406	Hs.285728	X95677	H.sapiens mRNA for ArgBPIB protein	2.9
	104577	Hs.132390	R71539	ESTs	2.9
	102772	Hs.161002	U83115	absent in melanoma 1	2.9

	131710	Hs.30985	AA233225	ESTs; Highly similar to (define not ava	2.9
	125231	Hs.268903	W84714	ESTs	2.9
	127380	Hs.15535	AI417137	Homo sapiens clone 24582 mRNA sequence	2.9
5	104229	Hs.61289	AB002346	inositol phosphate 5'-phosphatase 2 (syn	2.9
	126600	Hs.191385	AA699949	ESTs	2.9
	125175	Hs.303030	W52355	EST	2.9
	103849	Hs.34578	AA187045	ESTs; Weakly similar to !!!! ALU SUBFAMI	2.9
	102126	Hs.78961	U14575	protein phosphatase 1; regulatory (inhib	2.9
10	124906	Hs.107815	R87647	ESTs	2.9
	131148	Hs.303125	C00038	ESTs	2.9
	123158	Hs.218329	AA488658	heat shock 70kD protein 1	2.9
	133667	Hs.75462	U72649	Human BTG2 (BTG2) mRNA; complete cds	2.9
	105182	Hs.18271	AA191014	ESTs; Weakly similar to Ydr372cp [S.cere	2.9
	133968	Hs.232068	D15050	Human mRNA for transcription factor AREB	2.9
15	117425	Hs.336901	N27154	ESTs	2.9
	111087	Hs.37637	N59645	ESTs	2.9
	129641	Hs.11805	N66066	ESTs	2.9
	128639	Hs.102897	N91246	ESTs	2.9
20	133209	Hs.79265	AA114183	ESTs; Moderately similar to glutamate py	2.9
	135154	Hs.267812	AA126433	sorting nexin 4	2.9
	126838	Hs.279609	AA858097	pigment epithelium-derived factor	2.9
	103803	Hs.106149	AA127696	ESTs	2.9
	102139	Hs.2128	U15932	dual specificity phosphatase 5	2.9
25	128104		AA971000	op67g11.s1 Soares_NFL_T_GBC_S1 Homo sapi	2.8
	127834	Hs.337631	AA761415	nz22d08.s1 NC1_CGAP_GCB1 Homo sapiens cD	2.8
	133101	Hs.180952	AA488230	ESTs	2.8
	127250	Hs.217916	AI023717	ESTs	2.8
	135063	Hs.93883	D10537	myelin protein zero (Charcot-Marie-Tooth	2.8
	126323	Hs.68644	N45014	yy80g06.r1 Soares_multiple_sclerosis_2Nb	2.8
30	121873	Hs.145696	AA426270	ESTs	2.8
	122090	Hs.98684	AA432141	ESTs	2.8
	118728	Hs.322645	N73705	ESTs	2.8
	135400	Hs.99915	M23263	androgen receptor (dihydrotestosterone r	2.8
35	125278	Hs.129998	W93523	ESTs	2.8
	124387	Hs.109019	N27637	ESTs	2.8
	124803	Hs.12186	R45480	cyclin K	2.8
	H45968	Hs.32149	H45968	ESTs	2.8
	104261	Hs.5409	AF008442	RNA polymerase I subunit	2.8
40	105366	Hs.282093	AA236356	ESTs	2.8
	106070	Hs.5957	AA417761	Homo sapiens clone 24416 mRNA sequence	2.8
	131356	Hs.25960	M13241	v-myc avian myelocytomatosis viral relat	2.8
	112009	Hs.26255	R42714	EST	2.8
	133199	Hs.250175	AA609773	Homo sapiens clone 23904 mRNA sequence	2.8
45	110379	Hs.33130	H44825	ESTs	2.8
	103890	Hs.72085	AA236843	ESTs; Weakly similar to unknown [S.cerev	2.8
	128152		R20353	yg20f10.r1 Soares infant brain 1NIB Homo	2.8
	107008	Hs.23740	AA598710	ESTs	2.8
	135243	Hs.97101	AA215333	ESTs	2.8
50	103058	Hs.184510	X57348	stratifin	2.8
	132020	Hs.293845	AA428990	ESTs	2.8
	116354	Hs.292566	AA504262	ESTs	2.8
	125867	Hs.12372	H98141	ESTs	2.8
	120603	Hs.98541	AA282787	ESTs; Highly similar to (define not ava	2.8
55	115119	Hs.46847	AA256524	Human DNA sequence from clone 30M3 on ch	2.8
	133865	Hs.170290	F09315	discs; large (Drosophila) homolog 5	2.8
	109415	Hs.110826	AA227219	Homo sapiens CAGF9 mRNA; partial cds	2.8
	128687	Hs.23767	Z38910	ESTs	2.8
	109984	Hs.10299	H09594	ESTs; Moderately similar to !!!! ALU SUB	2.8
60	133179	Hs.66731	U81599	homeo box B13	2.8
	115998	Hs.336629	AA448488	ESTs; Weakly similar to zinc finger prot	2.8
	112180	Hs.25067	R49116	EST	2.8
	120428	Hs.173694	AA236822	ESTs; Moderately similar to (define not	2.8
	106241	Hs.6019	AA430108	ESTs	2.8
65	131060	Hs.22564	AA160890	myosin VI	2.8
	111383	Hs.40919	N94527	ESTs	2.8
	102123	Hs.1594	U14518	centromere protein A (17kD)	2.8
	102722	Hs.79981	U79242	Human clone 23560 mRNA sequence	2.8
	129887	Hs.274324	W92041	PCAF associated factor 65 alpha	2.8
	126663	Hs.181297	AA714635	ESTs	2.8

	104367	Hs.134342	H17438	ESTs; Weakly similar to septentrionalis	2.8
	107316	Hs.193700	T63174	ESTs; Moderately similar to Hs.193700	2.8
	128059	Hs.145096	AA972446	ESTs	2.8
	124447		N48000	ESTs	2.8
5	111398	Hs.125565	R00086	deafness; X-linked 1; progressive	2.8
	134085	Hs.79018	U20979	chromatin assembly factor I (150 kDa)	2.8
	124788	Hs.100912	R43543	ESTs	2.8
	112248	Hs.326416	R51361	ESTs	2.8
	121309	Hs.97312	AA402482	ESTs	2.8
10	103076	Hs.75319	X59618	ribonucleotide reductase M2 polypeptide	2.8
	107071	Hs.35198	AA609053	ESTs	2.8
	104425	Hs.35380	H88496	ESTs	2.8
	132991	Hs.62245	AA446906	solute carrier family 25 (mitochondrial)	2.8
	104968	Hs.29669	AA084602	ESTs	2.8
15	121153	Hs.97694	AA399640	ESTs	2.8
	131216	Hs.243901	D31058	ESTs	2.8
	109682	Hs.22869	F09299	ESTs	2.8
	131990	Hs.168818	H77734	ESTs; Moderately similar to roundabout 1	2.8
	132027	Hs.181444	N78844	ESTs; Weakly similar to R12C12.6 [C.eleg	2.8
20	127383	Hs.190478	AA447990	ESTs	2.8
	132598	Hs.530	M81379	collagen; type IV; alpha 3 (Goodpasture	2.8
	101121	Hs.1313	L09753	tumor necrosis factor (ligand) superfam	2.8
	123000	Hs.105640	AA479347	ESTs	2.8
	121329	Hs.1755	AA404324	ESTs	2.8
25	100481	Hs.121489	HG1098-HT1098	Cystatin D	2.7
	113803	Hs.283683	W42789	ESTs	2.7
	110934	Hs.169001	N48708	ESTs; Weakly similar to cytochrome P-450	2.7
	432888		T86823	ESTs	2.7
	121802	Hs.188898	AA424328	ESTs	2.7
30	130396	Hs.155313	AB002331	Human mRNA for KIAA0333 gene; partial cd	2.7
	121103	Hs.97697	AA398936	ESTs; Weakly similar to (define not ava	2.7
	131129	Hs.23240	R27296	ESTs	2.7
	130943	Hs.272429	D50855	calcium-sensing receptor (hypocalciuric	2.7
	134676	Hs.87819	W28051	ESTs; Weakly similar to keratin 9; cytos	2.7
35	111900	Hs.25318	R39044	ESTs	2.7
	106025	Hs.173334	AA412063	ESTs	2.7
	126144	Hs.40639	N39696	yx92a07.r1 Soares melanocyte 2NbHM Homo	2.7
	103248	Hs.75262	X77383	cathepsin O	2.7
	127230	Hs.274170	H30501	Homo sapiens Opa-interacting protein OIP	2.7
40	101584	Hs.84072	M35252	transmembrane 4 superfamily member 3	2.7
	124131	Hs.167489	H19980	ESTs	2.7
	129689	Hs.77873	AA130156	ESTs	2.7
	132892	Hs.9973	W92797	ESTs	2.7
	120827	Hs.132967	AA347717	ESTs	2.7
45	134579	Hs.85963	N23222	ESTs; Moderately similar to Hs.193700	2.7
	106149	Hs.256301	AA424881	ESTs	2.7
	132037	Hs.332541	AA203649	ESTs; Weakly similar to HEM45 [H.sapiens	2.7
	130542	Hs.179825	U64675	Human sperm membrane protein BS-63 mRNA,	2.7
	122851	Hs.99598	AA463627	ESTs	2.7
50	134983	Hs.196384	D28235	prostaglandin-endoperoxide synthase 2 (p	2.7
	120537	Hs.160422	AA262790	ESTs	2.7
	131036	Hs.174140	X64330	ATP citrate lyase	2.7
	133889	Hs.211582	AA099391	ESTs	2.7
	128847	Hs.106529	AA424199	zv81e01.r1 Soares_total_fetus_Nb2HF8_9w	2.7
55	112755	Hs.306044	R93802	ESTs	2.7
	423239		AA323591	EST26392 Cerebellum II Homo sapiens cDNA	2.7
	105031	Hs.12321	AA127240	ESTs	2.7
	126021	Hs.187516	AA775894	ESTs	2.7
	102116		U13706	Human ELAV-like neuronal protein 1 isofo	2.7
60	133394	Hs.237225	R16759	ESTs; Weakly similar to (define not ava	2.7
	104267	Hs.278439	C00358	ESTs	2.7
	107614	Hs.40241	AA004878	ESTs; Highly similar to (define not ava	2.7
	129809	Hs.1259	X55283	asialoglycoprotein receptor 2	2.7
	112109	Hs.283309	R45221	ESTs; Weakly similar to Hs.193700	2.7
65	128422		T85681	yd60c06.r1 Soares fetal liver spleen 1NF	2.7
	109494	Hs.43899	AA233702	ESTs	2.7
	118696	Hs.292284	N72086	Homo sapiens RNA polymerase III largest	2.7
	106053	Hs.36727	AA416963	ESTs; Highly similar to histone H2A [H.s	2.7
	104440	Hs.284380	L20492	gamma-glutamyltransferase 1	2.7

	129426	Hs.111323	AA412087	EST; Highly similar to (define not avai	2.7
	123798		AA620411	small inducible cytokine A5 (RANTES)	2.7
	106716	Hs.238928	AA464962	ESTs	2.7
5	103663		Z78291	Z78291 Homo sapiens brain fetus Homo sap	2.7
	114162	Hs.22265	Z38909	ESTs	2.7
	113063	Hs.5027	T32438	ESTs	2.7
	127897		AA773857	af80c09.r1 Soares_NhHMPu_S1 Homo sapiens	2.7
	130621	Hs.16803	AA621718	ESTs; Weakly similar to (define not ava	2.7
10	116245	Hs.42796	AA479958	ESTs; Highly similar to (define not ava	2.7
	125499		R11878	yf49d11.r1 Soares infant brain 1NIB Homo	2.7
	133960	Hs.77899	M19267	tropomyosin 1 (alpha)	2.7
	104470	Hs.246358	N28843	ESTs; Weakly similar to Similar to colla	2.7
	134982	Hs.92308	N46086	ESTs	2.7
	106803	Hs.284295	AA479114	ESTs	2.7
15	104899	Hs.285574	AA054726	ESTs	2.7
	125401	Hs.337585	AI204637	ESTs; Moderately similar to KIAA0350 [H.	2.7
	111253	Hs.15768	N70042	ESTs; Moderately similar to !!!! ALU SUB	2.7
	118449	Hs.164478	N66413	ESTs; Weakly similar to (define not ava	2.7
	134507	Hs.84318	M63488	replication protein A1 (70kD)	2.7
20	121609	Hs.98185	AA416867	EST	2.7
	113835	Hs.27475	W56590	ESTs	2.7
	113962	Hs.285290	W86375	ESTs; Highly similar to (define not ava	2.7
	121913	Hs.98558	AA428062	ESTs	2.7
25	108194	Hs.216717	AA057250	ESTs	2.7
	130799	Hs.12696	AA464273	ESTs	2.7
	123184	Hs.18166	AA489072	Homo sapiens mRNA for KIAA0870 protein;	2.7
	103420	Hs.173497	X97065	SEC23-like protein B	2.7
	106186	Hs.6315	AA427398	acetylserotonin N-methyltransferase-like	2.7
	101349		L77559	Homo sapiens DGS-B partial mRNA	2.7
30	112954	Hs.6655	T16559	ESTs	2.7
	133054	Hs.291079	R07876	ESTs; Weakly similar to unknown [S.cerev	2.7
	128131	Hs.25640	AI283162	claudin 3	2.6
	101864	Hs.75777	M95787	transgelin	2.6
35	111948	Hs.26303	R40752	ESTs	2.6
	130145	Hs.151051	U07620	protein kinase mitogen-activated 10 (MAP	2.6
	126507	Hs.23964	AI362218	ESTs	2.6
	117903	Hs.47111	N50740	ESTs	2.6
	116345	Hs.199067	AA496981	ESTs	2.6
40	132227	Hs.4248	AA412620	ESTs	2.6
	125746	Hs.274256	H03574	yj42b06.r1 Soares placenta Nb2HP Homo sa	2.6
	105073	Hs.89463	AA137034	ESTs	2.6
	102764		U82310	Homo sapiens unknown protein mRNA, parti	2.6
	131367	Hs.173933	AA456687	ESTs	2.6
45	130792	Hs.19500	AA307896	nuclear localization signal deleted in v	2.6
	107427	Hs.46736	W26975	ESTs	2.6
	117477	Hs.44175	N30328	ESTs	2.6
	106290	Hs.16364	AA435542	ESTs	2.6
	126829	Hs.7910	R11547	ESTs	2.6
50	118836	Hs.173001	N79820	ESTs	2.6
	100147	Hs.136348	D13666	osteoblast specific factor 2 (fasciclin	2.6
	104278	Hs.109253	C02582	ESTs; Highly similar to (define not ava	2.6
	135051	Hs.83484	C15324	ESTs	2.6
	126081	Hs.227835	AI346024	collagen; type I; alpha 1	2.6
55	123579		AA608983	af5d4.s1 Soares_testis_NHT Homo sapiens	2.6
	130115	Hs.149923	M31627	X-box binding protein 1	2.6
	101434	Hs.1430	M20218	coagulation factor XI (plasma thrombopla	2.6
	122962	Hs.104720	AA478429	ESTs; Moderately similar to !!!! ALU SUB	2.6
	126151	Hs.40808	AA324743	ESTs	2.6
60	128925	Hs.21851	D61676	Homo sapiens mRNA; cDNA DKFZp586J2118 (f	2.6
	128919	Hs.103391	L27559	insulin-like growth factor binding prote	2.6
	130296	Hs.154103	R09286	LIM protein (similar to rat protein kina	2.6
	128402	Hs.191637	AA457244	ESTs	2.6
	129273	Hs.109968	W63783	ESTs	2.6
	125483	Hs.7788	F07759	ESTs	2.6
65	132953	Hs.321264	AA029927	ESTs	2.6
	130963	Hs.21639	U57099	nuclear protein; marker for differentiat	2.6
	120614	Hs.194154	AA284281	ESTs; Weakly similar to !!!! ALU SUBFAM1	2.6
	123251	Hs.103267	AA490858	ESTs; Moderately similar to Rabin3 [R.no	2.6
	121710	Hs.96744	AA419011	ESTs	2.6

	125428	Hs.851	W74608	ESTs; Highly similar to (define not ava	2.6
	115906	Hs.82302	AA436616	ESTs	2.6
	108432		AA076626	Homo sapiens clone 23851 mRNA sequence	2.6
5	126191	Hs.191911	H97728	ESTs	2.6
	106164	Hs.281434	AA425773	ESTs	2.6
	111519	Hs.268615	R08165	ESTs	2.6
	134590	Hs.173840	W58612	ESTs	2.6
	102565		U59748	Human desert hedgehog (hDHH) mRNA, parti	2.6
10	129879	Hs.13109	AA194973	ESTs	2.6
	114264	Hs.334609	Z40074	ESTs	2.6
	106236	Hs.21104	AA429951	ESTs	2.6
	135192	Hs.321709	AF000234	purinergic receptor P2X; ligand-gated io	2.6
	109833	Hs.29889	H00580	ESTs	2.6
	105756	Hs.8535	AA303088	ESTs; Weakly similar to transformation-r	2.6
15	121422	Hs.97967	AA406210	ESTs	2.6
	130417	Hs.155485	U58522	Human huntingtin interacting protein (HI	2.6
	124312	Hs.102329	H94647	ESTs	2.6
	108998	Hs.97199	AA156058	ESTs	2.6
20	127081	Hs.180591	R88362	ESTs; Weakly similar to weak similarity	2.6
	129574	Hs.11463	AA458603	ESTs; Weakly similar to (define not ava	2.6
	112410	Hs.26904	R61680	ESTs	2.6
	123929	Hs.112981	AA621364	ESTs	2.6
	122905	Hs.104835	AA470070	ESTs	2.6
25	116399	Hs.110637	AA599729	Homo sapiens homeobox protein A10 (HOXA1	2.6
	130279	Hs.153934	AA424044	core-binding factor; runt domain; alpha	2.6
	130021	Hs.1435	M24470	guanosine monophosphate reductase	2.6
	100585	Hs.199160	HG2367-HT2463	Trithorax Homolog Hrx	2.6
	104965	Hs.30177	AA084104	ESTs	2.6
30	117711	Hs.46485	N45201	EST	2.6
	124792	Hs.48712	R44357	ESTs	2.6
	111299	Hs.74313	N73808	ESTs	2.6
	103616	Hs.32971	Z46973	phosphoinositide-3-kinase; class 3	2.6
	133629	Hs.195614	D13642	KIAA0017 gene product	2.6
35	126484	Hs.169977	A1086782	ESTs	2.6
	100858		HG4245-HT4515	Forkhead Family Afx1	2.6
	133547	Hs.301927	X02883	T-cell receptor; alpha (V;D;J;C)	2.6
	126680	Hs.133865	F07097	ESTs	2.6
	125739	Hs.92137	AA428557	v-myc avian myelocytomatosis viral oncog	2.6
40	102276	Hs.10247	U30999	Human (memc) mRNA, 3'UTR	2.6
	105586	Hs.191538	AA279137	ESTs	2.6
	103978	Hs.34136	AA307443	ESTs	2.6
	125054	Hs.268601	T80622	ESTs; Weakly similar to (define not ava	2.6
	114212	Hs.21201	Z39338	ESTs; Highly similar to (define not ava	2.6
	116959	Hs.40022	H79310	EST	2.6
45	109228	Hs.306995	AA193366	ESTs	2.6
	133989	Hs.78202	U29175	SWI/SNF related; matrix associated; acti	2.6
	100640	Hs.182183	HG2743-HT2845	Caldesmon 1, Alt. Splice 3, Non-Muscle	2.6
	133093	Hs.285996	AA598749	ESTs	2.6
	114306	Hs.6540	Z40861	ESTs	2.6
50	106060	Hs.171391	AA417287	C-terminal binding protein 2	2.5
	107748	Hs.60772	AA017258	EST	2.5
	100134	Hs.49	D13264	macrophage scavenger receptor 1	2.5
	133969	Hs.78	U13044	GA-binding protein transcription factor;	2.5
55	130992	Hs.74316	AA455001	ESTs	2.5
	127493	Hs.291701	AA808081	oc39a08.s1 NCI_CGAP_GCB1 Homo sapiens cD	2.5
	132869	Hs.203961	N26855	ESTs	2.5
	117570	Hs.44583	N34415	EST	2.5
	124644	Hs.109654	N91279	ESTs	2.5
60	103558	Hs.2785	Z19574	keratin 17	2.5
	132883	Hs.5897	AA047151	ESTs	2.5
	102009	Hs.82643	U02680	protein tyrosine kinase 9	2.5
	116058	Hs.20159	AA454156	ESTs	2.5
	121989	Hs.193784	AA430044	ESTs	2.5
	131257	Hs.24908	AA256042	ESTs	2.5
65	100320	Hs.75275	D50916	homolog of yeast (S. cerevisiae) ufd2	2.5
	102959	Hs.121524	X15722	glutathione reductase	2.5
	132969	Hs.6166	AA047616	ESTs	2.5
	130869	Hs.2057	AA128100	uridine monophosphate synthetase (orotat	2.5
	129645	Hs.118131	L38928	5;10-methenyltetrahydrofolate synthetase	2.5

	126399	Hs.83883	AA128075	zl16d08.r1 Soares_pregnant_uterus_NbHPU	2.5
	134069	Hs.78935	U29607	Homo sapiens eIF-2-associated p67 homolo	2.5
	109816	Hs.61960	F11013	ESTs; Weakly similar to KIAA0176 [H.sapi	2.5
5	134801	Hs.89695	X02160	insulin receptor	2.5
	104232	Hs.10587	AB002351	Human mRNA for KIAA0353 gene; partial cd	2.5
	107361	Hs.159486	U72513	Human RPL13-2 pseudogene mRNA; complete	2.5
	106057	Hs.289074	AA417067	ESTs	2.5
	134252	Hs.80720	AA031782	Homo sapiens mRNA; cDNA DKFZp586B1722 (f	2.5
10	128062	Hs.105547	AA379500	ESTs	2.5
	110009	Hs.6614	H10933	ESTs	2.5
	111375	Hs.20432	N93696	ESTs	2.5
	122642	Hs.99361	AA454186	ESTs	2.5
	127999	Hs.69851	AA837495	ESTs; Weakly similar to Wiskott-Aldrich	2.5
15	105029	Hs.13268	AA126855	ESTs	2.5
	105082	Hs.26765	AA143763	ESTs; Weakly similar to Similarity to S.	2.5

TABLE 1A show the accession numbers for those primekeys lacking unigeneID's for Table 1. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number
 CAT number: Gene cluster number
 Accession: Genbank accession numbers

Pkey	CAT number	Accessions
108552	111555_1	AA071210 AA069899 AA071438 AA084912 AA084803 AA079371 AA079370
126023	1596090_1	H57661 H58881
126086	1606216_1	H75681 H70975
102565	32479_1	AB010994 U59748 AA064660
101964	48158_7	S81578
125499	1562851_1	H10543 R11878
125596	1708455_1	R25698 R56582 R56018
118417	37186_1	AF080229 AF080231 AF080230 AF080232 AF080233 AF080234 BE550633 AI636743 AW614951 BE467547 AI680833 AI633818 N29986 U87592 U87593 U87590 U87591 S46404 U87587 AA463992 AW206802 AI970376 AI583718 AI672574 N25695 AW665466 AI818326 AA126128 AI480345 AW013827 AA248638 AI214968 AA204735 AA207155 AA206262 AA204833 AW003247 AW496808 AI080480 AI631703 AI651023 AI867418 AW818140 AA502500 AI206199 AI671282 AI352545 BE501030 AI652535 BE465762 AA206331 AW451866 AA471088 AA206342 AA204834 AA206100 AW021661 AA332922 N66048 AA703396 H92278 AW139734 H92683 U87589 U87595 H69001 U87594 BE466420 AI624817 BE466611 AI206344 AA574397 AA348354 AI493192 AA491830 R50173 R55192 R50320 AI732306 AI732305 AI820727 AI820728 R55191 R50319 R50227
125661	327827_1	H41694 H45213
125957	1583542_1	R98091 W92898
125982	1766315_1	AA364195 AA325029 AW962050
127248	227560_1	AA070545 AA131490 AA131373
103731	112052_1	AA330501 AA661567
127261	231687_1	AA331503 AA332751 AW962542
127265	232391_1	T16245 R19694 F13545 H10299 T66048 T65279 H18006
126659	1541209_1	AF116622 AI114507 AA640834 AA377999
127315	37938_1	AA130614 AA071410
103806	112618_1	AA906093 AA971000
128104	502608_1	H47610 R86920
104602	524482_2	F07973 R20353 AA442660
128152	297868_1	T77794 T85681
128422	1811283_1	AA773681 AA773857
127897	446527_1	BE298210 AI672315 AW086489 BE298417 AA455921 AA902537 BE327124 R14963 AA085210 AW274273 AI333584 AI369742 AI039658 AI885095 AI476470 AI287650 AI885299 AI985381 AW592624 AW340136 AI266556 AA456390 AI310815 AA484951
106566	120358_1	AI950087 N70208 R97040 N36809 AI308119 AW967677 N35320 AI251473 H59397 AW971573 R97278 W01059 AW967671 AA908598 AA251875 AI820501 AI820532 W87891 T85904 U71456 T82391 BE328571 T75102 R34725 AA884922 BE328517 AI219788 AA884444 N92578 F13493 AA927794 AI560251 AW874068 AL134043 AW235363 AA663345 AW008282 AA488964 AA283144 AI890387 AI950344 AI741346 AI689062 AA282915 AW102898 AI872193 AI763273 AW173586 AW150329 AI653832 AI762688 AA988777 AA488892 AI356394 AW103813 AI539642 AA642789 AA856975 AW505512 AI961530 AW629970 BE612881 AW276997 AW513601 AW512843 AA044209 AW856538 AA180009 AA337499 AW961101 AA251669 AA251874 AI819225 AW205862 AI683338 AI858509 AW276905 AI633006 AA972584 AA908741 AW072629 AW513996 AA293273 AA969759 N75628 N22388 H84729 H60052 T92487 AI022058 AA780419 AA551005 W80701 AW613456 AI373032 AI564269 F00531 H83488 W37181 W78802 R66056 AI002839 R67840 AA300207 AW959581 T63226 F04005
123147	219802_2	AA487961
130529	158447_1	AA178953 AA192740
123579	genbank_AA608983	AA608983
109175	genbank_AA180496	AA180496
100789	tigr_HT4163	S67998
100858	tigr_HT4515	U10072

	123798	579959_1	AA620411 AA287491
	102116	entrez_U13706	U13706
	102398	entrez_U42359	U42359
5	102764	entrez_U82310	U82310
	118475	genbank_N66845	N66845
	104776	genbank_AA026349	AA026349
	104787	genbank_AA027317	AA027317
	113702	genbank_T97307	T97307
10	113938	genbank_W81598	W81598
	122635	genbank_AA454085	AA454085
	108407	genbank_AA075519	AA075519
	108432	genbank_AA076626	AA076626
	108555	genbank_AA084963	AA084963
15	101349	entrez_L77559	L77559
	124447	genbank_N48000	N48000
	119071	genbank_R31180	R31180
	103520	entrez_Y10511	Y10511
	103663	genbank_Z78291	Z78291
20	128046	877605_1	AA873285 AI025762
	126959	546044_1	AA199853 AA206355
	123465	genbank_AA599033	AA599033

MISSING AT THE TIME OF PUBLICATION

TABLE 2: shows a preferred subset of the Accession numbers for genes found in Table 1 which are differentially expressed in prostate tumor tissue compared to normal prostate tissue.

5

Pkey: Unique Eos probeset identifier number
 ExAccn: Exemplar Accession number, Genbank accession number
 UnigeneID: Unigene number
 Unigene Title: Unigene gene title
 R1: Ratio of tumor to normal body tissue (Relaxed ratio (87/70))

10

15

20

25

30

35

40

45

50

55

60

65

Pkey	ExAccn	UnigeneID	Unigene Title	R1
131919	AA121266	Hs.272458	ESTs	37.2
120328	AA196979	Hs.290905	ESTs; Weakly similar to (define not ava	32.6
101486	M24902	Hs.1852	acid phosphatase; prostate	25.2
119073	R32894	Hs.279477	ESTs	24.8
133428	M34376	Hs.183752	microseminoprotein; beta-	23.8
128180	AA595348	Hs.171995	kallikrein 3; (prostate specific antigen	21.4
104080	AA402971	Hs.57771	Homo sapiens mRNA for serine protease (T	18.9
127537	AA569531	Hs.162859	ESTs	18.6
131665	R22139	Hs.30343	ESTs	17.4
101050	K01911	Hs.1832	neuropeptide Y	17.3
130771	N48056	Hs.1915	folate hydrolase (prostate-specific memb	17
107485	W63793	Hs.262476	S-adenosylmethionine decarboxylase 1	16.7
106155	AA425309	Hs.33287	ESTs	16.5
129534	R73640	Hs.11260	ESTs	16.4
100569	HG2261-HT2351		Antigen, Prostate Specific, Alt. Splice	16
101889	S39329	Hs.181350	kallikrein 2; prostatic	15.4
135389	U05237	Hs.99872	fetal Alzheimer antigen	15
133944	AA045870	Hs.7780	ESTs	12.5
130974	X57985	Hs.2178	H2B histone family; member Q	11.8
114768	AA149007	Hs.182339	ESTs	11.8
104660	AA007160	Hs.14846	ESTs	11.4
131061	N64328	Hs.268744	ESTs; Moderately similar to KIAA0273 [H.	10.9
126645	AI167942	Hs.61635	Homo sapiens BAC clone RG041D11 from 7q2	10.7
135153	N40141	Hs.95420	Homo sapiens mRNA for JM27 protein; comp	10.6
107033	AA599629	Hs.113314	ESTs	10.6
118417	N68048		ESTs; Weakly similar to polymerase [H.sa	10.5
126758	W37145	Hs.293960	ESTs	10.2
107102	AA609723	Hs.30652	ESTs	10.1
116787	H28581	Hs.15641	ESTs	10.1
115719	AA416997	Hs.59622	ESTs	10
123209	AA489711	Hs.203270	ESTs	9.9
101664	M60752	Hs.121017	H2A histone family; member A	9.8
112971	T17185	Hs.83883	ESTs	9.7
117984	N51919	Hs.106778	ESTs	9.7
129523	M30894	Hs.274509	T-cell receptor; gamma cluster	9.4
132964	AA031360	Hs.167133	ESTs	9.2
121853	AA425887	Hs.98502	ESTs	9
119617	W47380	Hs.55999	ESTs	8.9
105627	AA281245	Hs.23317	ESTs	8.8
101461	M22430	Hs.76422	phospholipase A2; group IIA (platelets;	8.7
124526	N62096	Hs.293185	yz61c5.s1 Soares_multiple_sclerosis_2NbH	8.5
133845	T68510	Hs.76704	ESTs	8.2
133354	AA055552	Hs.334762	ESTs; Weakly similar to KIAA0319 [H.sapi	8.1
119018	N95796	Hs.278695	ESTs	8
100394	D84276	Hs.66052	CD38 antigen (p45)	8
106579	AA456135	Hs.23023	ESTs	7.6
114965	AA250737	Hs.72472	ESTs	7.4
112033	R43162	Hs.22627	ESTs	7.1
102398	U42359		Human N33 protein form 1 (N33) gene, exo	7
101201	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin;	6.9
101803	M86546	Hs.155691	pre-B-cell leukemia transcription factor	6.8
120562	AA280036	Hs.302267	ESTs; Weakly similar to W01A6.c [C.elega	6.8

	109112	AA169379	Hs.257924	ESTs	6.8
	109795	F10707	Hs.326416	ESTs	6.7
	130336	X07730	Hs.171995	kallikrein 3; (prostate specific antigen	6.6
	131425	AA219134	Hs.26691	ESTs	6.6
5	132902	AA490969	Hs.59838	ESTs	6.6
	133724	U07919	Hs.75746	aldehyde dehydrogenase 6	6.5
	120215	Z41050	Hs.108787	Homo sapiens Mcd4p homolog mRNA; complet	6.5
	131881	AA010163	Hs.3383	upstream regulatory element binding prot	6.5
10	100727	X07290	Hs.334786	Human HF.12 gene mRNA	6.3
	121770	AA421714	Hs.278428	Homo sapiens mRNA for KIAA0896 protein;	6.3
	123475	AA599267	Hs.250526	ESTs; Weakly similar to ANKYRIN; BRAIN V	6.3
	133061	AB000584	Hs.296638	prostate differentiation factor	6.3
	116429	AA609710	Hs.279923	ESTs; Weakly similar to similar to GTP-b	6.2
15	101233	L29008	Hs.878	sorbitol dehydrogenase	6.2
	104691	AA011176	Hs.37744	ESTs	6.2
	127248	AA325029	EST27953	Cerebellum II Homo sapiens cDNA	6.2
	105500	AA256485	Hs.222399	ESTs	6.1
	130828	AA053400	Hs.203213	ESTs	5.9
	115357	AA281793	Hs.72988	ESTs	5.8
20	116334	AA491457	Hs.48948	ESTs	5.7
	120132	Z38839	Hs.125019	ESTs; Weakly similar to !!!! ALU SUBFAM I	5.6
	106375	AA443993	Hs.289072	ESTs	5.6
	124777	R41933	Hs.140237	ESTs; Weakly similar to neuronal thread	5.6
	101791	M83822	Hs.62354	Human beige-like protein (BGL) mRNA; par	5.5
25	117698	N41002	Hs.45107	ESTs	5.5
	122041	AA431407	Hs.98732	Homo sapiens Chromosome 16 BAC clone CIT	5.5
	133723	AA088851	Hs.262476	S-adenosylmethionine decarboxylase 1	5.5
	113938	W81598	ESTs		5.4
	133015	AA047036	Hs.246315	ESTs	5.4
30	108186	AA056482	Hs.7780	ESTs	5.3
	104466	N25110	Hs.326392	Human guanine nucleotide exchange factor	5.3
	104033	AA365031	Hs.98944	ESTs	5.3
	110844	N31952	Hs.167531	ESTs; Weakly similar to (define not ava	5.3
	129056	H70627	Hs.108336	ESTs; Weakly similar to !!!! ALU SUBFAM I	5.3
35	133493	AA284143	Hs.194369	Homo sapiens chromosome 1 atrophin-1 rel	5.3
	129184	W26769	Hs.109201	ESTs; Highly similar to (define not ava	5.2
	101448	M21389	Hs.195850	keratin 5 (epidermolysis bullosa simplex	5.1
	116188	AA464728	Hs.184598	ESTs; Weakly similar to !!!! ALU SUBFAM I	5.1
	105921	AA402613	Hs.169119	ESTs	5.1
40	103375	X91868	Hs.54416	sine oculis homeobox (Drosophila) homolo	5.1
	128871	AA400271	Hs.106778	ESTs; Highly similar to (define not ava	5.1
	116238	AA479382	Hs.47144	ESTs	5
	102913	X07696	Hs.80342	keratin 15	5
	103011	X52541	Hs.326035	early growth response 1	5
45	118981	N93839	Hs.39288	ESTs; Weakly similar to !!!! ALU SUBFAM I	5

TABLE 2A shows the accession numbers for those primekeys lacking unigeneID's for Table 2. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number
 CAT number: Gene cluster number
 Accession: Genbank accession numbers

Pkey	CAT number	Accession
118417	37186_1	AF080229 AF080231 AF080230 AF080232 AF080233 AF080234 BE550633 AI636743 AW614951 BE467547 AI680833 AI633818 N29986 U87592 U87593 U87590 U87591 S46404 U87587 AA463992 AW206802 AI970376 AI583718 AI672574 N25695 AW665466 AI818326 AA126128 AI480345 AW013827 AA248638 AI214968 AA204735 AA207155 AA206262 AA204833 AW003247 AW496808 AI080480 AI631703 AI651023 AI867418 AW818140 AA502500 AI206199 AI671282 AI352545 BE501030 AI652535 BE465762 AA206331 AW451866 AA471088 AA206342 AA204834 AA206100 AW021661 AA332922 N66048 AA703396 H92278 AW139734 H92683 U87589 U87595 H69001 U87594 BE466420 AI624817 BE466611 AI206344 AA574397 AA348354 AI493192
127248	227560_1	AA364195 AA325029 AW962050
107033	235652_1	AI141999 AA730176 R44544 R41778 AW300793 AW966157 AA918501 AA599629 AI082195 AI198537 AW006520 AW236663 AW151420 AI826987 AI810832 AI669102 AI201981 N27331 AA335566 T84622 BE085347 BE085269
102398	entrez_U42359	U42359
113938	genbank_W81598	W81598

TABLE 3: shows genes, including expression sequence tags, differentially expressed in prostate tumor tissue compared to normal tissue as analyzed using the Affymetrix/Eos Hu02 GeneChip array. Shown are the relative amounts of each gene expressed in prostate tumor samples and various normal tissue samples showing the highest expression of the gene.

Pkey: Unique Eos probeset identifier number
 ExAccn: Exemplar Accession number, Genbank accession number
 UnigenelD: Unigene number
 Unigene Title: Unigene gene title
 R1: Ratio of tumor to normal body tissue

Pkey	ExAccn	UnigenelD	Unigene Title	R1	
100131	D12485	Hs.11951	phosphodiesterase I/nucleotide pyrophosph	6.3	
100235	D29954	Hs.13421	KIAA0056 protein	5.1	
100570	HG2261-HT2352	Hs.171995	Antigen, Prostate Specific, Alt. Splice	9	
100819	HG4020-HT4290	Hs.2387	Transglutaminase	10.5	
101063	L00354	Hs.80247	cholecystokinin	8.5	
101247	L33801	Hs.78802	glycogen synthase kinase 3 beta	4.7	
101416	M17254	Hs.279477	v-ets avian erythroblastosis virus E26 o	4.7	
101447	M21305		Human alpha satellite and satellite 3 ju	11	
101485	M24736	Hs.89546	selectin E (endothelial adhesion molecucl	9.8	
101514	M28214	Hs.123072	RAB3B; member RAS oncogene family	6.2	
101626	M57399	Hs.44	pleiotrophin (heparin binding growth fac	8.4	
101663	M60750	Hs.2178	H2B histone family; member A	4.9	
101758	M77836	Hs.79217	pyrroline-5-carboxylate reductase 1	5.4	
101768	M81118	Hs.78989		7.5	
101817	M88163	Hs.152292	SWI/SNF related; matrix associated; acti	5.5	
101888	M99701	Hs.95243	transcription elongation factor A (SII)-	5.7	
102031	U04898	Hs.2156	RAR-related orphan receptor A	13.2	
102052	U07559	Hs.505	ISL1 transcription factor; LIM/homeodoma	8.9	
102221	U24576	Hs.3844	LIM domain only 4	5.6	
102233	U26173	Hs.79334	nuclear factor; interleukin 3 regulated	7.4	
102302	U33052	Hs.69171	protein kinase C-like 2	8.2	
102348	U37519	Hs.87539	aldehyde dehydrogenase 8	5.9	
102457	U48807	Hs.2359	dual specificity phosphatase 4	5.1	
102473	U49957	Hs.180398	LIM domain-containing preferred transloc	5.7	
102669	U71207	Hs.29279	eyes absent (Drosophila) homolog 2	9	
102698	U75272	Hs.1867	progastricsin (pepsinogen C)	10.6	
102751	U80034	Hs.68583	mitochondrial intermediate peptidase	15.6	
102823	U90914	Hs.5057	carboxypeptidase D	4.9	
102869	X02544	Hs.572	orosomucoid 1	22.6	
103031	X54667	Hs.123114	cystatin S	4.7	
103043	X55733	Hs.93379	eukaryotic translation initiation factor	4.9	
103093	X60708	Hs.44926	dipeptidylpeptidase IV (CD26; adenosine	5.8	
103376	X92098	Hs.323378	coated vesicle membrane protein	5.2	
103401	X95240	Hs.54431	specific granule protein (28 kDa); cyste	7.4	
103613	Z46629	Hs.2316	SRY (sex-determining region Y)-box 9 (ca	5.2	
103677	Z83806		H.sapiens mRNA for axonemal dynein heavy	4.9	
103962	AA298180	Hs.83243	ESTs	6	
104084	AA410529	Hs.30732	ESTs	6.4	
104257	AF006265	Hs.9222	estrogen receptor-binding fragment-assoc	6.8	
104301	D45332	Hs.6783	ESTs	10.5	
104769	AA025887	Hs.293943	ESTs; Weakly similar to !!!! ALU SUBFAMI	6.3	
104851	AA040882	Hs.10290	U5 snRNP-specific 40 kDa protein (hPrp8-	4.9	
104896	AA054228	Hs.23165	ESTs	5.8	
104956	AA074880	Hs.20509	ESTs; Weakly similar to hypothetical pro	6.4	
104957	AA074919	Hs.10026	ESTs; Weakly similar to ORF YJL063c [S.c	4.8	
104967	AA084506	Hs.291000	ESTs	6.5	
105099	AA150776	Hs.23729	Homo sapiens clone 24405 mRNA sequence	7	
105298	AA233459	Hs.26369	ESTs	5.1	

	105304	AA233553	Hs.190325	ESTs	4.7
	105370	AA236476	Hs.22791	ESTs; Weakly similar to transmembrane pr	10.3
	105427	AA251330	Hs.28248	ESTs	5
5	105542	AA261858	Hs.266957	ESTs; Weakly similar to heat shock prote	8.8
	105628	AA281251	Hs.79828	ESTs; Weakly similar to putative zinc fi	5.5
	105640	AA281623	Hs.6685	ESTs; Weakly similar to KIAA0742 protein	8
	105645	AA282138	Hs.11325	ESTs	14
	105691	AA287097	Hs.289068	transcription factor 4	6.3
10	105730	AA292701	Hs.5364	DKFZP564I052 protein	4.9
	105808	AA393808	Hs.286131	KIAA0438 gene product	7
	105826	AA398243	Hs.194477	ESTs; Moderately similar to similar to N	5
	105903	AA401433	Hs.200016	ESTs; Weakly similar to diphosphoinosito	9.9
	105906	AA401633	Hs.22380	ESTs	11.5
15	106065	AA417558	Hs.25206	ESTs	5.1
	106094	AA419461	Hs.23317	ESTs	10.9
	106157	AA425367	Hs.34892	ESTs	6.6
	106184	AA426643	Hs.10762	ESTs	8.5
	106211	AA428240	Hs.126083	ESTs	8.4
20	106213	AA428258	Hs.8769	Homo sapiens mRNA; cDNA DKFZp564E153 (fr	5.7
	106272	AA432074	Hs.323099	ESTs	5.8
	106369	AA443828	Hs.288856	ESTs	6.3
	106400	AA447621	Hs.94109	ESTs	5.4
	106474	AA450212	Hs.42484	Homo sapiens mRNA; cDNA DKFZp564C053 (fr	9.2
25	106507	AA452584	Hs.267819	protein phosphatase 1; regulatory (inhib	5.6
	106523	AA453441	Hs.31511	ESTs	4.7
	106532	AA453628	Hs.37443	ESTs	4.7
	106557	AA455087	Hs.22247	ESTs	5.7
	106575	AA456039	Hs.105421	ESTs	7.2
30	106618	AA459249	Hs.8715	ESTs; Weakly similar to Similarity with	5.6
	106820	AA481037	Hs.12592	ESTs	5.4
	106846	AA485223	Hs.34892	ESTs	5.3
	106973	AA505141	Hs.11923	Human DNA sequence from clone 167A19 on	7.5
	107110	AA609952	Hs.12784	KIAA0293 protein	6.1
	107127	AA620504	Hs.179898	ESTs	7.1
35	107159	AA621340	Hs.10600	ESTs; Weakly similar to ORF YKR081c [S.c	5.2
	107217	D51095	Hs.35861	DKFZP586E1621 protein	15.1
	107365	U78294	Hs.111256	arachidonate 15-lipoxygenase; second typ	4.7
	107630	AA007218	Hs.60178	ESTs	5.3
40	107734	AA016225	Hs.7517	ESTs	4.8
	107760	AA018042	Hs.252085	EST	7.6
	107997	AA037388	Hs.82223	Human DNA sequence from clone 141H5 on c	10.5
	108012	AA039616	Hs.173334	ESTs	6.5
	108520	AA084138	Hs.46786	ESTs	7.9
45	108583	AA088276	Hs.68826	ESTs	5.6
	108613	AA100967	Hs.69165	ESTs	6
	108664	AA113349	Hs.69588	EST	6.3
	108677	AA115629	Hs.118531	ESTs	5.9
	108807	AA129968	Hs.49376	ESTs; Weakly similar to PROTEIN PHOSPHAT	5.8
50	108910	AA136590		ESTs	5
	108933	AA147224	Hs.337232	ESTs	12.7
	108948	AA149579	Hs.118258	ESTs	6.8
	109014	AA156790	Hs.262036	ESTs	15.3
	109124	AA171529	Hs.183887	ESTs	6.1
55	109142	AA176438	Hs.41295	ESTs	5.1
	109277	AA196332	Hs.86043	ESTs	5.5
	109342	AA213620		Homo sapiens mRNA; cDNA DKFZp586M1418 (f6	
	109562	F01811	Hs.187931	ESTs; Moderately similar to voltage-gate	10.8
	109565	F01930	Hs.23648	ESTs	7
60	109648	F04600	Hs.7154	ESTs	9.9
	109799	F10770	Hs.180378	Homo sapiens clone 669 unknown mRNA; com	6.4
	109859	H02308	Hs.20792	ESTs	5.3
	110181	H20276	Hs.31742	ESTs	16.8
	110854	N32919	Hs.27931	ESTs	10
65	110924	N47938	Hs.12940	yy84a09.s1 Soares_multiple_sclerosis_2Nb	5.6
	111046	N55514	Hs.318584	ESTs	6.9
	111091	N59858	Hs.33032	Homo sapiens mRNA; cDNA DKFZp434N185 (fr	5.2
	111157	N66613	Hs.99364	ESTs	5
	111164	N66857	Hs.122489	ESTs; Weakly similar to !!!! ALU CLASS C	5.6
	111221	N68869	Hs.15119	ESTs	6.2

	111348	N90041	Hs.9585	ESTs	5.4
	111353	N90430	Hs.6616	ESTs	5.3
	111495	R07210	Hs.9683	ESTs	5.8
	111540	R08850	Hs.9786	ESTs	6
5	111579	R10657	Hs.167115	KIAA0830 protein	12.6
	111581	R10684	Hs.5794	ESTs	7.1
	111734	R25375	Hs.128749	ESTs	6.2
	111861	R37460	Hs.25231	ESTs	9.4
10	111870	R37778	Hs.18685	ESTs; Weakly similar to hypothetical pro	6.5
	111937	R40431	Hs.14846	Homo sapiens mRNA; cDNA DKFZp564D016 (fr	4.8
	111987	R42036	Hs.6763	KIAA0942 protein	6.4
	112184	R49173	Hs.330242	ESTs	5.6
	112286	R53765	Hs.158135	KIAA0981 protein	9.3
15	112380	R59740	Hs.5740	ESTs	4.7
	112452	R63841	Hs.157461	ESTs	6
	112601	R79111	Hs.78225	annexin A1	5.4
	112753	R93696	Hs.169882	ESTs	5.8
	112902	T09262	Hs.129190	ESTs	5.1
	112984	T23457	Hs.289014	ESTs	4.9
20	113021	T23855	Hs.129836	KIAA1028 protein	10.8
	113083	T40530	Hs.266957	ESTs; Weakly similar to heat shock prote	5.7
	113200	T57773	Hs.10263	ESTs	7.3
	113494	T88878	Hs.86538	ESTs	8.7
25	113849	W60439	Hs.8858	ESTs; Moderately similar to cbp146 [M.mu	4.9
	113883	W72382	Hs.11958	oxidative 3 alpha hydroxysteroid dehydro	4.7
	113950	W85765	Hs.30504	Homo sapiens mRNA; cDNA DKFZp434E082 (fr	6.7
	113986	W87462	Hs.21894	ESTs	5.9
	113989	W87544	Hs.268828	ESTs	4.7
30	114124	Z38595	Hs.125019	ESTs; Highly similar to KIAA0886 protein	21.3
	114340	Z41395	Hs.143611	ESTs	9.6
	114346	Z41450	Hs.130489	ESTs	5.2
	114435	AA018216	Hs.164975	Bicaudal D (Drosophila) homolog 1	7.4
	114463	AA025370	Hs.40109	KIAA0872 protein	8.2
	114652	AA101416	Hs.107149	ESTs; Weakly similar to PTB-ASSOCIATED S	5.4
35	114721	AA131450	Hs.103822	ESTs	4.8
	114730	AA133527	Hs.331328	ESTs; Weakly similar to The KIAA0138 gen	5.1
	114833	AA234362	Hs.87159	ESTs; Moderately similar to CGI-66 prote	5.5
	114860	AA235112	Hs.42179	ESTs; Moderately similar to similar to m	6.3
	114884	AA235811	Hs.293672	ESTs	5.2
40	114895	AA236177	Hs.76591	KIAA0887 protein	4.7
	114908	AA236545	Hs.54973	ESTs	5.2
	114932	AA242751	Hs.16218	KIAA0903 protein	5.7
	115084	AA255566	Hs.42484	Homo sapiens mRNA; cDNA DKFZp564C053 (fr	5.2
45	115140	AA258030	Hs.279938	ESTs; Weakly similar to supported by GEN	5.9
	115468	AA287061	Hs.48499	ESTs; Highly similar to Bdeight protein	4.7
	115583	AA398913	Hs.45231	LDOC1 protein	7.6
	115709	AA412519	Hs.58279	ESTs	4.8
	115772	AA423972	Hs.131740	ESTs	5
50	115774	AA424029	Hs.288390	ESTs; Moderately similar to dynamin; int	5.4
	115776	AA424038	Hs.81897	ESTs	5
	115821	AA427528	Hs.130965	ESTs; Weakly similar to ZINC FINGER PROT	13.7
	115955	AA446121	Hs.44198	Homo sapiens BAC clone RG054D04 from 7q3	10.6
	116024	AA451748	Hs.83883	Human DNA sequence from clone 718J7 on c	6.8
55	116108	AA457566	Hs.28777	ESTs	6
	116117	AA459117	Hs.31575	SEC63; endoplasmic reticulum translocon	7.3
	116146	AA460701	Hs.15423	ESTs	5.5
	116296	AA489033	Hs.62601	Homo sapiens mRNA; cDNA DKFZp586K1318 (f	5.7
	116379	AA521472	Hs.71252	ESTs	5.9
	116393	AA599463	Hs.306051	protein phosphatase 2 (formerly 2A); reg	5.9
60	116401	AA599963	Hs.59698	ESTs	7.9
	116416	AA609219	Hs.39982	ESTs	9.2
	116587	D59325	Hs.121429	ESTs	5.2
	116601	D80055	Hs.45140	ESTs	4.9
	116684	F09156	Hs.66095	ESTs	7.2
65	116722	F13654		HSFIH32 Stratagene cat#937212 (1992) Hom	5.5
	116766	H13260	Hs.95097	ESTs	5.9
	117453	N29568	Hs.108319	thyroid hormone receptor-associated prot	6.9
	117557	N33920	Hs.44532	diubiquitin	4.8
	117708	N45114	Hs.126280	ESTs	6.3

	118001	N52151	Hs.47447	ESTs	11.4
	118229	N62339	Hs.166254	heat shock 90kD protein 1; alpha	6.2
	118599	N69207	Hs.203697	ESTs	5.8
5	118645	N70358	Hs.125180	growth hormone receptor	7.1
	118873	N89881	Hs.44577	ESTs	6
	118985	N94303	Hs.55028	ESTs	9.3
	119107	R42424	Hs.63841	ESTs	6
	119126	R45175	Hs.117183	ESTs	17.9
10	119271	T16387	Hs.65328	ESTs	6
	119367	T78324	Hs.250895	ESTs	5
	119721	W69440	Hs.48376	ESTs	15.4
	119741	W70205	Hs.43670	kinesin family member 3A	10.1
	119780	W72967	Hs.191381	ESTs; Weakly similar to hypothetical pro	5.3
15	120217	Z41078	Hs.66035	ESTs	4.8
	120266	AA173939	Hs.205442	ESTs; Weakly similar to inner centromere	8.8
	120294	AA190888	Hs.153881	ESTs; Highly similar to NY-REN-62 antigen	4.9
	120418	AA236010	Hs.26613	Homo sapiens mRNA; cDNA DKFZp586F1323 (f	4.7
	120486	AA253400	Hs.137569	tumor protein 63 kDa with strong homolog	5.6
20	120524	AA261852	Hs.192905	ESTs	4.9
	120571	AA280738	Hs.34892	ESTs	8.8
	120596	AA282074	Hs.237323	ESTs	6.2
	120713	AA292655	Hs.96557	ESTs	9.9
	120992	AA398246	Hs.97594	ESTs	16.4
25	121429	AA406293	Hs.41167	ESTs	6.9
	121503	AA412049	Hs.290347	ESTs	7.6
	121512	AA412105	Hs.193736	ESTs	5.8
	121816	AA424814	Hs.48827	ESTs	4.6
	122027	AA431302	Hs.98721	EST; Weakly similar to N-copine [H.sapie	5.6
	122294	AA437311	Hs.98927	ESTs	5.7
30	122411	AA446859	Hs.99083	ESTs	6.5
	122791	AA460158	Hs.129836	KIAA1028 protein	12.4
	122792	AA460225	Hs.99519	ESTs	5.1
	122969	AA478539	Hs.104336	ESTs	4.9
35	123095	AA485724	Hs.27413	ESTs	5.4
	123100	AA485957	Hs.306219	Homo sapiens clone 25032 mRNA sequence	5
	123295	AA495981	Hs.250830	ESTs	4.7
	123311	AA496252	Hs.105069	ESTs	7.4
	123583	AA609006	Hs.111240	ESTs	9.1
40	123619	AA609200		ESTs	4.7
	123645	AA609310	Hs.188691	ESTs	4.8
	123709	AA609651	Hs.112742	ESTs	7
	123968	C14333	Hs.108327	damage-specific DNA binding protein 1 (1	5
	124178	H45996	Hs.97101	putative G protein-coupled receptor	6.8
45	124352	N21626	Hs.102406	ESTs	10.2
	124357	N22401		yw37g07.s1 Morton Fetal Cochlea Homo sap	10.6
	124515	N58172	Hs.109370	ESTs	14.2
	124911	R88992	Hs.174195	ESTs	4.8
	125154	W38419		ESTs	4.7
50	125992	W01626		za36e07.r1 Soares fetal liver spleen 1NF	5.1
	126802	AA947601	Hs.97056	ESTs	5.1
	126812	Z36290	Hs.173933	ESTs; Weakly similar to NUCLEAR FACTOR 1	4.6
	127080	AA662913	Hs.190173	ESTs	5
	127308	AA507628	Hs.334390	ESTs	4.8
55	127370	AI024352	Hs.70337	immunoglobulin superfamily; member 4	4.7
	127386	AI457411	Hs.106728	ESTs	4.8
	127965	AA828760	Hs.292059	ESTs	4.8
	128172	AI400862	Hs.265130	ESTs	5
	128305	AI039722	Hs.279009	ESTs	5.8
60	128420	AI088155	Hs.41296	ESTs; Weakly similar to unknown [H.sapie	17
	128467	AA176446	Hs.180428	ESTs; Weakly similar to hypothetical 43.	4.8
	128610	L38608	Hs.10247	activated leucocyte cell adhesion molecu	7.9
	128625	AA242816	Hs.102652	ESTs; Weakly similar to KIAA0437 [H.sapi	8.1
	128651	AA446990	Hs.103135	ESTs	6.5
65	129088	AA215971	Hs.194431	KIAA0992 protein	5.2
	129136	N26391	Hs.250723	ESTs	5.1
	129171	AA234048	Hs.7753	calumenin	5.8
	129229	AA211941	Hs.109643	polyadenylate binding protein-interactin	5.8
	129386	N27524	Hs.260024	Cdc42 effector protein 3	5.2
	129467	AA410311	Hs.44208	ESTs	5.1

	129564	H22136	Hs.75295	guanylate cyclase 1; soluble; alpha 3	16.3
	129699	AA458578	Hs.12017	KIAA0439 protein; homolog of yeast ubiq	9.2
	129821	F11019	Hs.12696	cortactin SH3 domain-binding protein	8.6
5	129823	X00948	Hs.105314	relaxin 2 (H2)	9.1
	129847	W46767	Hs.296178	ESTs; Weakly similar to RNA POLYMERASE I	5.4
	129912	AA047344	Hs.107213	ESTs; Highly similar to NY-REN-6 antigen	6.5
	129958	L20591	Hs.1378	annexin A3	5.1
	129977	J04076	Hs.1395	early growth response 2 (Krox-20 (Drosop	8.6
10	130061	U82256	Hs.172851	arginase; type II	7.4
	130241	U78313	Hs.153203	MyoD family inhibitor	4.9
	130466	N21679	Hs.180059	ESTs	5.8
	130541	X05608	Hs.211584	neurofilament; light polypeptide (68kD)	6.7
	130619	AA477739	Hs.12532	ESTs	6.4
	130925	N71935	Hs.169378	multiple PDZ domain protein	7.9
15	130938	AA013250	Hs.21398	ESTs; Moderately similar to PUTATIVE GLU	6.2
	130971	H20332	Hs.301444	signal sequence receptor; gamma (translo	6.4
	131066	F09006	Hs.22588	ESTs	5
	131126	F09012	Hs.181326	myotubularin related protein 2	6.4
	131310	J02960	Hs.2551	adrenergic; beta-2-; receptor; surface	7.9
20	131487	AA253220	Hs.27373	Homo sapiens mRNA; cDNA DKFZp564O1763 (f	5.9
	131561	X59841	Hs.294101	pre-B-cell leukemia transcription factor	7.6
	131562	U90551	Hs.28777	H2A histone family; member L	5.1
	131579	N62922	Hs.29088	ESTs	11
	131629	AA442119	Hs.238809	ESTs	4.9
25	131682	AA428368	Hs.30654	ESTs	4.8
	131699	R68657	Hs.90421	ESTs; Moderately similar to !!!! ALU SUB	6.5
	131795	N32724	Hs.32317	Sox-like transcriptional factor	5.6
	132053	H93381	Hs.38085	ESTs; Weakly similar to putative glycine	7.2
	132122	U65092	Hs.40403	Cbp/p300-interacting transactivator; wit	5.6
30	132191	AA449431	Hs.288361	KIAA0741 gene product	8
	132256	AA608856	Hs.431	murine leukemia viral (bmi-1) oncogene h	5.5
	132482	AA429478	Hs.238126	ESTs; Highly similar to CGI-49 protein [6.6
	132533	AA021608	Hs.172510	ESTs	5.8
	132572	AA448297	Hs.237825	signal recognition particle 72kD	6.2
35	132581	R42266	Hs.52256	ESTs; Weakly similar to beta-TrCP protei	16
	132700	N47109	Hs.5521	ESTs	6.8
	132701	AA279359	Hs.55220	BCL2-associated athanogene 2	5.3
	132725	L41887	Hs.184167	splicing factor; arginine/serine-rich 7	7.8
40	132783	N74897	Hs.278894	DEAD/H (Asp-Glu-Ala-Asp/His) box polypep	5.9
	132790	X75535	Hs.168670	peroxisomal farnesylated protein	8
	132939	U76189	Hs.61152	exostoses (multiple)-like 2	5.2
	133142	F03321	Hs.65874	ESTs	5.2
	133342	U29589	Hs.7138	cholinergic receptor; muscarinic 3	10.3
	133434	AA278852	Hs.30212	ESTs	5.8
45	133453	M68941	Hs.73826	protein tyrosine phosphatase; non-recept	4.9
	133520	X74331	Hs.74519	primase; polypeptide 2A (58kD)	13.1
	133544	T33873	Hs.74624	protein tyrosine phosphatase; receptor t	4.6
	133608	D13315	Hs.75207	glyoxalase I	4.8
	133626	H75939	Hs.75277	Homo sapiens mRNA; cDNA DKFZp586M141 (fr	5
50	133633	D21262	Hs.75337	nucleolar phosphoprotein p130	6.3
	133797	S66431	Hs.76272	retinoblastoma-binding protein 2	6
	133928	N34096	Hs.7766	ubiquitin-conjugating enzyme E2E 1 (homo	5.4
	134095	U47414	Hs.79069	cyclin G2	5.2
	134249	N89827	Hs.80667	RALBP1 associated Eps domain containing	6.5
55	134321	AA418230	Hs.8172	ESTs	7
	134453	X70683	Hs.83484	SRY (sex determining region Y)-box 4	4.7
	134542	X57025	Hs.85112	insulin-like growth factor 1 (somatomedi	7.7
	134570	U68615	Hs.172280	SWI/SNF related; matrix associated; acti	6.4
	134592	U82613	Hs.289104	Alu-binding protein with zinc finger dom	5.4
60	134654	W23625	Hs.8739	ESTs; Weakly similar to ORF YGR200c [S.c	5
	134666	AA482319	Hs.8752	putative type II membrane protein	5.4
	134806	Z49099	Hs.89718	spermine synthase	6.7
	134951	AA431480	Hs.169358	ESTs	9.8
	135066	X04602	Hs.93913	interleukin 6 (interferon; beta 2)	5.7
65	135155	AA358268	Hs.166556	ESTs; Moderately similar to transcriptio	4.9
	135411	L10333	Hs.99947	reticulon 1	5.3
	300023	M10098		AFFX control: 18S ribosomal RNA	4.6
	300254	AW079607	Hs.55610	ESTs; Weakly similar to ZnT-3 [H.sapiens	7.8
	300273	AW013907	Hs.167531	ESTs; Moderately similar to predicted us	11.5

	300319	AW157646	Hs.153506	ESTs; Weakly similar to microtubule-acti	8.5
	300566	H86709	Hs.326392	son of sevenless (Drosophila) homolog 1	5.8
	300578	AI989417	Hs.134289	ESTs	4.4
	300671	AI239706	Hs.93810	ESTs	7.9
5	300675	AA039352	Hs.125034	ESTs; Weakly similar to ORF YDL040c [S.c	4.5
	300680	AW468066	Hs.24817	ESTs; Weakly similar to KIAA0986 protein	5.2
	300762	AI497778	Hs.20509	ESTs	6.4
	300810	AI076890	Hs.146847	ESTs	5.8
	300813	AA406411	Hs.208341	ESTs; Weakly similar to KIAA0989 protein	10.6
10	300823	AI863068	Hs.106823	ESTs; Weakly similar to putative zinc fi	5.6
	300834	AF109300	Hs.147924	ESTs	6.7
	300923	AW136372	Hs.1852	ESTs	7.6
	300962	AA593373	Hs.293744	ESTs	5.5
	301015	AA947682	Hs.20252	ESTs; Weakly similar to Chain A; Cdc42hs	7
15	301042	AI659131	Hs.197733	ESTs	24.9
	301242	AW161535	Hs.23782	ESTs	11.8
	301254	AI049624	Hs.283390	EST cluster (not in UniGene) with exon h	4.3
	301262	H29500	Hs.7130	ESTs; Moderately similar to N-copine [H.	4.3
20	301388	AA156879	Hs.262036	ESTs; Weakly similar to ZINC FINGER PROT	6.6
	301563	AI802946	Hs.44208	ESTs; Weakly similar to match to ESTs AA	5.7
	301656	AW008475	Hs.151258	EST cluster (not in UniGene) with exon h	6.8
	301689	Z44810	Hs.301789	ESTs; Weakly similar to similar to C.ele	6.3
	301783	AL046347	Hs.83937	Homo sapiens PAC clone DJ1159004 from 7p	6.2
25	301805	AI800004	Hs.142846	ESTs; Weakly similar to MesP1 [M.musculu	8.5
	301846	R20002	Hs.6823	ESTs; Weakly similar to intrinsic factor	4.6
	301891	AF131855	Hs.279591	Homo sapiens clone 25056 mRNA sequence	6.3
	302005	AI869666	Hs.123119	ESTs	36.8
	302056	AI457532	Hs.30488	ESTs; Moderately similar to ROSA26AS [M.	9.5
30	302067	H05698	Hs.222399	ESTs; Weakly similar to protein-tyrosine	5.8
	302099	AL021397	Hs.137576	ribosomal protein L34 pseudogene 1	8.8
	302147	AB022660	Hs.151717	KIAA0437 protein	5.9
	302214	AJ001454	Hs.159425	Homo sapiens mRNA for testican-3	4.3
	302236	AI128606	Hs.6557	zinc finger protein 161	4.3
35	302358	D81150	Hs.322848	EST cluster (not in UniGene) with exon h	5.5
	302410	NM_004917	Hs.218366	EST cluster (not in UniGene) with exon h	26.8
	302486	AC003682	Hs.183512	multiple UniGene matches	8.2
	302582	NM_000522	Hs.249195	EST cluster (not in UniGene) with exon h	6.4
	302785	AA425562	Hs.11065	EST cluster (not in UniGene) with exon h	5
40	302792	AA343696	Hs.46821	ESTs; Weakly similar to putative [H.sapi	4.8
	302881	AA508353	Hs.105314	relaxin 1 (H1)	78.8
	302892	N58545	Hs.42346	histone deacetylase 3	8.5
	302970	AW118352	Hs.312679	EST cluster (not in UniGene) with exon h	7.4
	302977	AW263124	Hs.315111	EST cluster (not in UniGene) with exon h	5.5
	303029	AF199613		EST cluster (not in UniGene) with exon h	4.6
45	303125	AF161352	Hs.111782	EST cluster (not in UniGene) with exon h	5.8
	303280	AI571580	Hs.170307	ESTs	4.3
	303306	AA215297	Hs.61441	EST cluster (not in UniGene) with exon h	6.4
	303309	AL134164	Hs.145416	ESTs	6.6
50	303344	AA255977	Hs.250646	ESTs; Highly similar to ubiquitin-conjug	19.5
	303380	AA298471	Hs.326567	EST cluster (not in UniGene) with exon h	6.6
	303401	AA758552	Hs.309497	ESTs	6.8
	303525	AW516519	Hs.273294	ESTs	4.8
	303526	AA348111	Hs.96900	ESTs	12.1
55	303540	AA355607	Hs.309490	ESTs; Weakly similar to MMSET type I [H.	8.2
	303572	AW338520	Hs.242540	ESTs	8.4
	303685	AW500106	Hs.23643	EST cluster (not in UniGene) with exon h	4.9
	303699	D30891	Hs.19525	EST cluster (not in UniGene) with exon h	15.7
	303702	AW500748	Hs.224961	ESTs; Weakly similar to 73 kDa subunit o	6.3
60	303718	AI741397	Hs.114658	ESTs	4.6
	303722	AA521510	Hs.145010	ESTs	12.5
	303732	AW502405	Hs.125759	ESTs; Weakly similar to tumor suppressor	4.3
	303735	AA707750	Hs.169055	ESTs; Weakly similar to cis-Golgi matrix	5.4
	303752	AI017286	Hs.5957	EST cluster (not in UniGene) with exon h	5.3
	303753	AW503733	Hs.9414	ESTs	13
65	303813	AI275850	Hs.114658	EST cluster (not in UniGene) with exon h	7.8
	304053	R00493	Hs.125565	translocase of inner mitochondrial membr	4.8
	304218	N66373	Hs.27973	ESTs; Weakly similar to ZK354.7 [C.elega	6
	305200	AA668128	Hs.45207	EST singleton (not in UniGene) with exon	5.7
	306716	AI024916	Hs.251354	ESTs	5.7

	307848	AI364186		EST singleton (not in UniGene) with exon	7.3
	307871	AI368665	Hs.31476	EST singleton (not in UniGene) with exon	5.4
	308050	AI460004	Hs.31608	EST singleton (not in UniGene) with exon	8.1
5	308362	AI613519	Hs.105749	EST singleton (not in UniGene) with exon	5.5
	308923	AI863051	Hs.279815	ESTs	4.4
	309116	AI927149	Hs.29797	ribosomal protein L10	4.5
	309375	AW075342	Hs.9271	EST singleton (not in UniGene) with exon	7.4
	309674	AW205604	Hs.266009	ESTs; Weakly similar to !!!! ALU SUBFAM1	5
10	310095	AI921750	Hs.144871	ESTs	5
	310098	AI685841	Hs.161354	ESTs	11.6
	310250	AI478629	Hs.158465	ESTs	5.8
	310365	AI262148	Hs.145569	ESTs	9.7
	310382	AI734009	Hs.127699	EST cluster (not in UniGene)	10.4
15	310409	AI612775	Hs.145710	ESTs	4.6
	310431	AI420227	Hs.149358	ESTs	72.9
	310573	AW292180	Hs.156142	ESTs	7.6
	310598	AI338013	Hs.140546	ESTs	9.2
	310639	AW269082	Hs.175162	ESTs	4.5
20	310787	AW262580	Hs.147674	ESTs	4.9
	310816	AI973051	Hs.224965	ESTs	7.6
	311251	AI655662	Hs.197698	ESTs	41.3
	311280	AI767957	Hs.198248	ESTs; Weakly similar to Y38A8.1 gene pro	4.5
	311330	AI679524	Hs.201629	ESTs; Moderately similar to !!!! ALU SUB	4.6
25	311515	AW136713	Hs.23862	ESTs	5.9
	311574	AI824863	Hs.211420	ESTs	4.8
	311587	AI828254	Hs.271019	ESTs	5.8
	311596	AI682088	Hs.79375	ESTs	26.4
	311631	AI809519	Hs.27133	ESTs	6.4
30	311688	AW025661	Hs.240090	ESTs	7.4
	311783	AI682478	Hs.13528	EST	4.6
	311826	AA765470	Hs.85092	ESTs	6.7
	311853	AW014013	Hs.107056	ESTs	5.3
	311901	R16890	Hs.137135	ESTs	5.6
35	311932	AW451654	Hs.257482	ESTs	4.3
	312153	AA759250	Hs.118625	cytochrome b-561	11
	312182	AA834800	Hs.326263	EST cluster (not in UniGene)	16.9
	312242	AI380207	Hs.125276	ESTs	4.7
	312296	C01367	Hs.127128	ESTs	5.3
40	312407	R46180	Hs.153485	ESTs	6.2
	312424	AA847398	Hs.291997	ESTs	4.8
	312425	R49353	Hs.293892	ESTs	5.2
	312480	R68651	Hs.144997	ESTs	9.5
	312518	C17785	Hs.182738	ESTs	6.3
45	312521	AA033609	Hs.239884	ESTs	11.2
	312527	AI695522	Hs.191271	ESTs	4.7
	312539	AI004377	Hs.200360	ESTs	7
	312546	AI623511	Hs.118567	ESTs	5.1
	312563	AA976064	Hs.180842	ESTs	6.5
50	312623	AA694607	Hs.176956	EST cluster (not in UniGene)	10.8
	312857	AA772279	Hs.126914	ESTs	5
	312890	AI813654	Hs.5957	ESTs	5.8
	312903	AA939266	Hs.278626	ESTs	7.7
	312905	H92571	Hs.234478	ESTs	6.5
55	312976	AA836271	Hs.125830	ESTs	4.6
	312983	AI079278	Hs.269899	ESTs	5.1
	312996	AA249018	Hs.154331	EST cluster (not in UniGene)	7
	313035	N36417	Hs.144928	ESTs	6.3
	313166	AI801098	Hs.151500	ESTs	4.3
60	313188	AI039702	Hs.179573	collagen; type I; alpha 2	4.8
	313218	AA827805	Hs.124296	ESTs	5
	313226	AI200281	Hs.123910	ESTs	5.9
	313325	AI420611	Hs.127832	ESTs	4.6
	313326	AI088120	Hs.122329	ESTs	7.4
65	313425	AA745689	Hs.186838	ESTs; Weakly similar to similar to zinc	6.3
	313499	AI261390	Hs.146085	ESTs	5.6
	313540	AI797301	Hs.5740	ESTs	5.9
	313568	AW467376	Hs.129640	ESTs	4.3
	313569	AI273419	Hs.135146	ESTs; Weakly similar to ZK1058.5 [C.eleg	4.6
	313603	AW468119	Hs.287631	EST cluster (not in UniGene)	6.8

	313615	AW295194	Hs.301997	DKFZP434N126 protein	5.2
	313625	AW468402	Hs.254020	ESTs	7.8
	313634	AA688292	Hs.337786	ESTs	4.4
	313635	AA507227	Hs.6390	ESTs	8.1
5	313638	AI753075	Hs.104627	ESTs	6.7
	313670	C16690	Hs.23767	EST cluster (not in UniGene)	4.4
	313671	W49823	Hs.104613	ESTs	4.4
	313676	AA861697	Hs.120591	EST cluster (not in UniGene)	13.4
10	313703	AI161293	Hs.280380	ESTs; Weakly similar to KIAA0525 protein	10
	313712	AA768553	Hs.74170	ESTs	5.2
	313800	AW296132	Hs.55098	ESTs	5.4
	313979	AI535895	Hs.221024	ESTs	4.3
	314121	AI732100	Hs.187619	ESTs	13.6
15	314123	AW245993	Hs.223394	ESTs	6.4
	314171	AI821895	Hs.193481	ESTs	29.4
	314188	AL138431	Hs.164243	ESTs	4.6
	314219	AL036001	Hs.48376	ESTs	5.7
	314236	AA743396	Hs.189023	ESTs	4.9
	314237	AA732359	Hs.96264	ESTs	4.4
20	314284	AA731431	Hs.293464	EST cluster (not in UniGene)	6.4
	314305	AI280112	Hs.125232	ESTs	5.3
	314343	AI754701	Hs.328476	ESTs; Weakly similar to alternatively sp	6.2
	314530	AI052358	Hs.193726	ESTs	4.5
25	314691	AW207206	Hs.136319	ESTs	17
	314695	AW502698	Hs.118152	ESTs	8.9
	314785	AI538226	Hs.32976	ESTs	9.4
	314801	AA481027	Hs.109045	ESTs; Weakly similar to ORF YGR245c [S.c	8
	314864	AA493811	Hs.294068	ESTs	6
	314907	AI672225	Hs.222886	ESTs	19.3
30	314916	AA548906	Hs.122244	ESTs	4.5
	314954	AA521381	Hs.187726	ESTs	5.3
	314981	AA524953	Hs.293334	ESTs	4.6
	315021	AA533447	Hs.312989	EST cluster (not in UniGene)	5.1
35	315051	AW292425	Hs.163484	EST	15.5
	315052	AA876910	Hs.134427	ESTs	20
	315073	AW452948	Hs.257631	ESTs	5.3
	315084	AI821085		ESTs	8.2
	315214	AI915927	Hs.34771	ESTs	5.4
40	315220	AI420753	Hs.66731	ESTs	5.1
	315278	AI985544	Hs.12450	ESTs	5.8
	315282	AI222165	Hs.144923	ESTs	4.5
	315368	AW291563	Hs.104696	ESTs	8
	315369	AA764918	Hs.256531	ESTs	4.8
	315378	AI263393	Hs.145008	ESTs	6.2
45	315379	AI378329	Hs.126629	ESTs	5.4
	315402	AW293424	Hs.75354	ESTs	5.1
	315442	AA977935	Hs.127274	ESTs	6.6
	315443	AW003416	Hs.160604	ESTs	5.5
	315528	R37257	Hs.184780	ESTs	8.1
50	315593	AW198103	Hs.158154	ESTs	9.9
	315634	AA837085	Hs.220585	ESTs	7.8
	315705	AW449285	Hs.313636	ESTs	8.9
	315707	AI418055	Hs.161160	ESTs	5.1
55	315714	AA744015	Hs.298138	EST cluster (not in UniGene)	6.1
	315740	T05558	Hs.156880	EST cluster (not in UniGene)	6.8
	315762	AI391470	Hs.158618	ESTs	5.3
	315769	AA744875	Hs.189413	ESTs	5
	315843	AA679430	Hs.191897	ESTs	5.7
	315990	AI800041	Hs.190555	ESTs	9.2
60	316012	AA764950	Hs.119898	ESTs	4.3
	316036	AA708016	Hs.190389	ESTs	5.9
	316055	AA693880	Hs.6947	EST cluster (not in UniGene)	6.7
	316074	AW517542	Hs.293273	ESTs	5.5
	316100	AW203986	Hs.213003	ESTs	5.1
65	316169	AI127483	Hs.120451	ESTs	8.2
	316442	AA760894	Hs.153023	ESTs	17.1
	316491	AA766025	Hs.186854	EST	4.6
	316504	AW135854	Hs.132458	ESTs	4.3
	316667	AW015940	Hs.232234	ESTs	7.6

	316854	AA831215	Hs.159066	ESTs; Weakly similar to predicted using	5.1
	316905	AW138241	Hs.210846	ESTs	6.4
	317008	AW051597	Hs.143707	ESTs	4.4
5	317019	AA864968	Hs.127699	ESTs	11
	317194	AW445167	Hs.126036	ESTs	13.5
	317224	D56760	Hs.93029	ESTs	8.7
	317404	AI806867	Hs.126594	ESTs	8.7
	317501	AA931245	Hs.137097	ESTs	11.1
	317548	AI654187	Hs.195704	ESTs	14.2
10	317651	AW292779	Hs.169799	ESTs	5.8
	317758	AI733277	Hs.128321	ESTs	5.4
	317850	N29974	Hs.152982	EST cluster (not in UniGene)	11.4
	317869	AW295184	Hs.129142	ESTs; Weakly similar to DEOXYRIBONUCLEAS	13.8
	317902	AI828602	Hs.211265	ESTs	5.3
15	317916	AI565071	Hs.159983	ESTs	7.7
	318239	AI085198	Hs.164226	ESTs	13.1
	318268	AI817736	Hs.182490	ESTs	6.2
	318327	AW294013	Hs.200942	ESTs	4.6
20	318363	R45530	Hs.1440	gamma-aminobutyric acid (GABA) A recepto	6
	318428	AI949409	Hs.194591	ESTs	12.3
	318464	AI151010	Hs.157774	ESTs	4.3
	318524	AW291511	Hs.159066	ESTs	25.9
	318540	T30280	Hs.274803	EST cluster (not in UniGene)	7
	318591	AW206806	Hs.115325	ESTs	4.8
25	318615	AI133617	Hs.10177	ESTs	5.5
	318646	AW175665	Hs.278695	ESTs	5.7
	318667	AI493742	Hs.165210	ESTs	11
	318668	W26276	Hs.136075	ESTs	5.9
	318753	AA578265	Hs.7130	copine IV	5.5
30	319080	Z45131	Hs.23023	ESTs	16.9
	319181	F06504	Hs.27384	EST cluster (not in UniGene)	4.6
	319191	AF071538	Hs.79414	prostate epithelium-specific Ets transcr	6.6
	319233	R21054	Hs.180532	ESTs	4.9
	319586	D78808	Hs.283683	ESTs	8.2
35	319750	AA621606	Hs.117956	ESTs	9.3
	319763	AA460775	Hs.6295	ESTs	14.3
	319824	AA424266	Hs.123642	EST cluster (not in UniGene)	12.8
	319838	AA337642	Hs.95262	nuclear factor related to kappa B bindin	5.1
	319913	AA179304	Hs.271586	ESTs; Moderately similar to !!!! ALU SUB	4.3
40	319964	T80579	Hs.290270	ESTs	5.8
	320076	AI653733	Hs.271593	ESTs	8.5
	320102	AW296219	Hs.115325	RAB7; member RAS oncogene family-like 1	9.8
	320187	T99949	Hs.303428	EST cluster (not in UniGene)	9.8
	320211	AL039402	Hs.125783	DEME-6 protein	7.9
45	320324	AF071202	Hs.139336	ATP-binding cassette; sub-family C (CFTR	56.2
	320455	R49889	Hs.24144	EST cluster (not in UniGene)	8.3
	320464	AI089817	Hs.237146	ESTs	5.4
	320561	NM_006953	Hs.159330	EST cluster (not in UniGene)	7
50	320574	AL049443	Hs.161283	Homo sapiens mRNA; cDNA DKFZp586N2020 (f	4.4
	320576	AL049977	Hs.162209	Homo sapiens mRNA; cDNA DKFZp564C122 (fr	6.7
	320654	AW263086	Hs.118112	ESTs	6
	320796	AF038966	Hs.31218	secretory carrier membrane protein 1	13.5
	320800	AI681006	Hs.71721	ESTs	6.2
	320813	AW360847	Hs.16578	ESTs	9.3
55	320853	AI473796	Hs.135904	ESTs	8.1
	320856	D59945	Hs.65366	EST cluster (not in UniGene)	6
	320899	AA633772	Hs.116796	ESTs	9.2
	320918	AW195012	Hs.293970	ESTs	5
	320973	H19732	Hs.247917	ESTs	5.9
60	321099	AA018386	Hs.64341	ESTs	4.6
	321190	H52462	Hs.163872	EST cluster (not in UniGene)	5.8
	321318	AB033041	Hs.137507	EST cluster (not in UniGene)	8.4
	321382	AW372449	Hs.175982	EST cluster (not in UniGene)	7.3
	321441	AW297633	Hs.118498	ESTs	14.7
65	321538	H80483	Hs.46903	EST cluster (not in UniGene)	9.2
	321609	H86021	Hs.182538	ESTs; Weakly similar to hMmTRA1b [H.sapi	4.8
	321636	AI791838	Hs.193465	ESTs	5.5
	321638	AI356352	Hs.108932	ESTs	4.6
	321644	AI204177	Hs.237396	ESTs	6.6

	321681	AA233821	Hs.190173	EST cluster (not in UniGene)	4.6
	321726	X91221	Hs.144465	EST cluster (not in UniGene)	5
	321758	U29112	Hs.196151	EST cluster (not in UniGene)	6.2
5	321877	AL109784	Hs.189222	EST cluster (not in UniGene)	4.6
	321899	N55158	Hs.29468	ESTs	4.6
	321902	AA746374	Hs.145010	ESTs	8.2
	322007	AW410646	Hs.164649	ESTs	5.1
	322055	AL137646	Hs.146001	EST cluster (not in UniGene)	4.3
10	322092	AF085833	Hs.135624	EST cluster (not in UniGene)	4.3
	322221	AI890619	Hs.179662	nucleosome assembly protein 1-like 1	4.4
	322278	AF086283		EST cluster (not in UniGene)	5.8
	322303	W07459	Hs.157601	EST cluster (not in UniGene)	22
	322437	AW393804	Hs.170253	ESTs; Weakly similar to rabaptin-4 [H.sa	4.4
15	322493	AF143235	Hs.279819	EST cluster (not in UniGene)	7.2
	322782	AA056060	Hs.202577	EST cluster (not in UniGene)	18.4
	322811	AA782292	Hs.105872	ESTs	6.9
	322818	AW043782	Hs.293616	ESTs	10.7
	322826	AI807883	Hs.180059	ESTs	5
20	322887	AI986306	Hs.86149	ESTs; Weakly similar to KIAA0969 protein	11.9
	322889	AA081924	Hs.124918	ESTs	7.1
	322924	AA669253	Hs.136075	ESTs	4.5
	322982	AI351191	Hs.128430	ESTs	6.6
	322994	AA422116	Hs.191461	ESTs	4.7
25	323040	AA336609	Hs.10862	ESTs	6.9
	323041	AL118747	Hs.26691	EST cluster (not in UniGene)	8.3
	323045	AA148950	Hs.188836	ESTs	4.6
	323048	AL118923	Hs.175110	EST cluster (not in UniGene)	7.5
	323070	AA157726	Hs.264330	ESTs	7.5
	323071	AA157867	Hs.5722	ESTs	4.7
30	323097	Z44354	Hs.296261	guanine nucleotide binding protein (G pr	4.9
	323131	AA176982	Hs.270124	EST cluster (not in UniGene)	6.1
	323136	AL120351	Hs.30177	EST cluster (not in UniGene)	4.3
	323175	AI827137	Hs.336454	ESTs	6.2
35	323218	AF131846	Hs.13396	Homo sapiens clone 25028 mRNA sequence	6.3
	323226	AF055019	Hs.21906	Homo sapiens clone 24670 mRNA sequence	12.6
	323236	AA363148	Hs.293960	ESTs	10.9
	323262	AI829770	Hs.190642	ESTs	7.6
	323276	AA836452	Hs.323822	ESTs	7.6
	323287	AA639902	Hs.104215	ESTs	24.7
40	323335	AI655499	Hs.161712	ESTs	14.1
	323341	AL134875	Hs.108646	ESTs	5.3
	323362	AL135067	Hs.117182	ESTs	6.1
	323486	C05278	Hs.299221	ESTs; Moderately similar to [PYRUVATE DE	8.5
45	323496	AI826801	Hs.300700	ESTs	4.5
	323507	H71721	Hs.128387	ESTs	4.4
	323545	AI814405	Hs.224569	ESTs	5.8
	323623	AA314280	Hs.146589	EST cluster (not in UniGene)	5
	323663	AW263526	Hs.243023	ESTs	7.7
50	323691	AA317561	Hs.145599	EST cluster (not in UniGene)	5.9
	323810	AA740405	Hs.108806	ESTs	6.2
	323846	AA337621	Hs.137635	ESTs	6
	323929	AA354940	Hs.145958	ESTs	10.7
	323959	AI636775	Hs.6831	ESTs	5.4
55	323996	AA367032	Hs.217882	ESTs	5.8
	323997	AA844907	Hs.274454	EST cluster (not in UniGene)	4.4
	324019	AW177009		EST cluster (not in UniGene)	4.6
	324130	AL046575	Hs.130198	ESTs	11
	324295	AI146686	Hs.143691	ESTs	13.7
60	324296	AI524039	Hs.192524	ESTs	6.8
	324307	AA627642	Hs.4994	transducer of ERBB2; 2 (TOB2)	4.9
	324330	AA884766		EST cluster (not in UniGene)	4.3
	324385	F28212	Hs.284247	EST cluster (not in UniGene)	4.7
	324430	AA464018	Hs.184598	EST cluster (not in UniGene)	13.6
65	324452	AW014022	Hs.170953	ESTs	7.6
	324547	AW501974	Hs.74170	ESTs	5.6
	324603	AW016378	Hs.292934	ESTs	24.2
	324617	AA508552	Hs.195839	ESTs	54
	324618	AI346282	Hs.87159	ESTs	4.6
	324620	AA448021	Hs.94109	EST cluster (not in UniGene)	5.7

	324626	AI685464	ESTs	9	
	324658	AI694767	Hs.129179 ESTs	22	
	324676	AW503943	Hs.112451 ESTs	4.9	
5	324691	AI217963	Hs.293341 ESTs; Weakly similar to Pro-a2(XI) [H.sa	10.6	
	324696	AA641092	Hs.257339 ESTs	10.2	
	324713	AW340249	Hs.163440 ESTs	5.5	
	324715	AI739168	Hs.131798 EST cluster (not in UniGene)	7.2	
	324718	AI557019	Hs.116467 ESTs	34.4	
10	324720	AA578904	Hs.292437 ESTs	4.8	
	324752	AI279919	Hs.272072 ESTs; Moderately similar to !!!! ALU SUB	7.9	
	324753	AA612626	Hs.144871 EST cluster (not in UniGene)	5.2	
	324790	AI334367	Hs.159337 ESTs	7.6	
	324801	AI819924	Hs.14553 ESTs	12.6	
	324804	AI692552	ESTs	6.5	
15	324845	AA361016	Hs.337533 ESTs	4.5	
	324888	AI564134	Hs.136102 KIAA0853 protein	4.4	
	324929	AI741633	Hs.125350 ESTs	6.5	
	324961	AA613792	EST cluster (not in UniGene)	5.1	
	325108	AA401863	Hs.22380 ESTs	7.1	
20	326816		CH.20_hs gi 6552458	9.6	
	326997		CH.21_hs gi 5867660	4.8	
	327098		CH.21_hs gi 6682516	4.3	
	328492		CH.07_hs gi 5868455	5.8	
	329362		CH.X_hs gi 5868837	4.3	
25	329929		CH.16_p2 gi 6165201	5.5	
	329960		CH.16_p2 gi 5091594	7.6	
	330020		CH.16_p2 gi 6671887	6	
	330211		CH.05_p2 gi 6013592	12.6	
	330384	M23263	androgen receptor (dihydrotestosterone r	9	
30	330430	HG2261-HT2352	Hs.321110	Antigen, Prostate Specific, Alt. Splice	13.8
	330546	U31382	Hs.299867 guanine nucleotide binding protein 4	6	
	330551	U39840	hepatocyte nuclear factor 3; alpha	4.9	
	330658	AA319514	Hs.30732 ESTs	6	
	330700	AA037415	Hs.20999 ESTs	5.5	
35	330704	AA056557	Hs.6759 ESTs	5.1	
	330705	AA102571	Hs.157078 ESTs	11.7	
	330706	AA121140	Hs.177576 ESTs; Moderately similar to kynurenine a	14.5	
	330712	AA167269	Hs.52620 ESTs	5	
	330725	AA252033	Hs.24052 ESTs; Weakly similar to !!!! ALU SUBFAM1	7.2	
40	330732	AA281092	Hs.35254 ESTs	4.9	
	330762	AA449677	Hs.15251 Human DNA sequence from clone 437M21 on	18.5	
	330763	AA450200	Hs.143187 FK506-binding protein 3 (25kD)	4.3	
	330772	AA479114	Hs.11356 ESTs	5.8	
	330786	D60374	EST	4.6	
45	330892	AA149579	Hs.91202 ESTs	15.3	
	330949	H01458	Hs.142896 ESTs	10.3	
	330977	H20826	Hs.315181 ESTs	4.4	
	331017	N24619	Hs.108920 ESTs	11.8	
	331099	R36671	Hs.14846 ESTs	11.6	
50	331128	R51361	Hs.268714 ESTs	4.8	
	331151	R82331	Hs.268838 ESTs	13	
	331195	T64447	Hs.168439 ESTs	4.9	
	331320	AA262999	Hs.300141 ESTs	4.8	
	331321	AA278355	Hs.87929 ESTs	6.1	
55	331337	AA287662	Hs.118630 ESTs	9.2	
	331348	AA400596	Hs.88143 ESTs	9.9	
	331359	AA416979	Hs.81897 ESTs	4.3	
	331383	AA454543	Hs.43543 ESTs	4.6	
	331422	F10802	Hs.237339 ESTs; Moderately similar to !!!! ALU SUB	4.9	
60	331442	H77381	Hs.41223 ESTs	7.5	
	331466	N21680	Hs.43455 ESTs	5.4	
	331479	N27154	Hs.44076 ESTs	6.5	
	331490	N32912	Hs.291039 ESTs; Weakly similar to hypothetical 43.	12.5	
	331493	N34357	Hs.93817 ESTs	4.6	
65	331561	N62780	Hs.48703 ESTs	9.2	
	331615	N92352	Hs.5472 ESTs	4.6	
	331659	W48868	Hs.334305 ESTs	8.7	
	331696	Z38907	Hs.65949 KIAA0888 protein	10.3	
	331811	AA404500	Hs.187958 ESTs	4.8	

	331848	AA417039	Hs.98268	signal recognition particle 72kD	7.5
	331873	AA429445	Hs.98640	ESTs	6.5
	331889	AA431407	Hs.98802	Homo sapiens Chromosome 16 BAC clone CIT	33.6
	331967	AA460158	Hs.99589	KIAA1028 protein	6.8
5	331974	AA464518	Hs.105322	ESTs	5.3
	332043	AA490831	Hs.201591	ESTs	10.8
	332076	AA599477	Hs.291156	ESTs	4.4
	332173	F09281	Hs.100725	ESTs	5.5
	332247	N58172		ESTs	14.2
10	332249	N62096	Hs.194140	ESTs	7.2
	332325	T79428	Hs.339667	ESTs	5.6
	332396	AA340504		ESTs; Weakly similar to similar to human	21.2
	332434	N75542	Hs.237731	transcription factor 4	15.3
	332493	N95495	Hs.56729	ESTs; Highly similar to GTP-binding prot	7.1
15	332522	L38503	Hs.178357	glutathione S-transferase theta 2	6.6
	332526	AA281753	Hs.17731	inositol 1;4;5-triphosphate receptor; ty	5.8
	332530	M31682	Hs.19280	inhibin; beta B (activin AB beta polypep	5.5
	332533	M99487	Hs.325825	folate hydrolase (prostate-specific memb	38.1
	332538	N48715	Hs.20991	ESTs	6.5
20	332546	D84454	Hs.22587	solute carrier family 35 (UDP-galactose	4.8
	332594	AA279313	Hs.32951	methyl CpG binding protein 2	5.6
	332610	AA412405	Hs.40513	ESTs; Weakly similar to BETA GALACTOSIDA	5.6
	332661	N95742	Hs.6390	ESTs	6.9
	332697	T94885	Hs.75725	carboxypeptidase E	24.3
25	332712	D26070	Hs.79306	inositol 1;4;5-triphosphate receptor; ty	9.9
	332716	L00058	Hs.79630	v-myc avian myelocytomatosis viral oncog	5.6
	332726	R72029	Hs.83428	synaptophysin-like protein	5
	332781	AA233258		ESTs; Weakly similar to D1007.5 [C.elega	4.5
	332797			CH22_FGENES.6_2	30.8
30	332798			CH22_FGENES.6_5	66.8
	332799			CH22_FGENES.6_6	19.8
	332933			CH22_FGENES.38_7	5.6
	332980			CH22_FGENES.54_1	5.5
	332984			CH22_FGENES.54_6	4.9
35	333168			CH22_FGENES.94_1	4.7
	333169			CH22_FGENES.94_2	4.4
	333452			CH22_FGENES.157_1	4.8
	333456			CH22_FGENES.157_5	4.3
	333458			CH22_FGENES.157_7	4.6
40	333611			CH22_FGENES.217_6	4.7
	333621			CH22_FGENES.219_5	5.5
	333814			CH22_FGENES.282_2	7.1
	333849			CH22_FGENES.290_8	6.2
	333949			CH22_FGENES.303_5	4.3
45	333951			CH22_FGENES.303_7	4.9
	333955			CH22_FGENES.303_11	5.6
	334150			CH22_FGENES.339_1	5.1
	334223			CH22_FGENES.360_4	20.3
	334297			CH22_FGENES.372_3	9.4
50	334443			CH22_FGENES.387_2	4.6
	334444			CH22_FGENES.387_4	5.6
	334447			CH22_FGENES.387_7	13.1
	334570			CH22_FGENES.405_11	5.4
	334749			CH22_FGENES.427_1	5.3
55	334777			CH22_FGENES.430_9	4.7
	334960			CH22_FGENES.465_29	5.2
	335179			CH22_FGENES.504_9	8.8
	335293			CH22_FGENES.527_6	4.7
	335550			CH22_FGENES.576_11	5.1
60	335581			CH22_FGENES.581_19	5.7
	335586			CH22_FGENES.581_25	4.3
	335809			CH22_FGENES.617_6	6.2
	335810			CH22_FGENES.617_7	5.8
	335822			CH22_FGENES.619_7	7.1
65	335824			CH22_FGENES.619_11	8.5
	335853			CH22_FGENES.626_5	4.3
	335886			CH22_FGENES.632_4	4.3
	336034			CH22_FGENES.678_5	6.8
	336441			CH22_FGENES.827_7	7.6

	336624	CH22_FGENES.6-3	43.3
	336625	CH22_FGENES.6-4	37.9
	336679	CH22_FGENES.43-7	5.3
	337577	CH22_C65E1.GENSCAN.8-1	4.9
5	338255	CH22_EM:AC005500.GENSCAN.276-3	13.4
	338260	CH22_EM:AC005500.GENSCAN.279-10	4.6
	338561	CH22_EM:AC005500.GENSCAN.421-5	4.6
	338562	CH22_EM:AC005500.GENSCAN.421-6	4.3
	338759	CH22_EM:AC005500.GENSCAN.517-6	5.1
10	338763	CH22_EM:AC005500.GENSCAN.517-16	5.5
	338764	CH22_EM:AC005500.GENSCAN.517-17	7.1

TABLE 3A shows the accession numbers for those primekeys lacking unigeneID's for Table 3. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

10 Pkey: Unique Eos probeset identifier number
 CAT number: Gene cluster number
 Accession: Genbank accession numbers

15	Pkey	CAT number	Accession
	123619	371681_1	AA602964 AA609200
	116722	143512_1	Z24878 AA494098 F13654 AA494040 AA143127
	103677	41847_1	Z83806 AJ132091 AJ132090
20	125992	1589048_1	H48372 W01626
	109342	genbank_AA213620	AA213620
	125154	genbank_W38419	W38419
	101447	entrez_M21305	M21305
	124357	genbank_N22401	N22401
25	108910	genbank_AA136590	AA136590
	322278	47271_1	W69304 AF086283 W69200
	315084	350959_1	AI821085 AW973464 AA554802 AI821831 AA657438 AA640756 AA650339
	324019	262792_1	AW177009 AI381610
	324330	300543_1	AA884766 AW974271 AA592975 AA447312
30	324626	336411_1	AI685464 AW971336 AA513587 AA525142
	303029	37699_1	AF199613 AF108756
	324804	398093_1	AI692552 AI393343 AI800510 AI377711 F24263 AA661876
	324961	376239_1	AA613792 AW182329 T05304 AW858385
	329362	c_x_hs	
35	336624	CH22_4071FG_6_3	
	336625	CH22_4072FG_6_4	
	336679	CH22_4157FG_43_7	
	338255	CH22_6856FG_LINK_EM:AC00	
	338260	CH22_6863FG_LINK_EM:AC00	
40	329929	c16_p2	
	329960	c16_p2	
	338561	CH22_7294FG_LINK_EM:AC00	
	338562	CH22_7295FG_LINK_EM:AC00	
	338759	CH22_7581FG_LINK_EM:AC00	
45	338763	CH22_7585FG_LINK_EM:AC00	
	338764	CH22_7586FG_LINK_EM:AC00	
	333168	CH22_400FG_94_1_LINK_EM:A	
	333169	CH22_401FG_94_2_LINK_EM:A	
	333452	CH22_702FG_157_1_LINK_EM:	
50	333456	CH22_706FG_157_5_LINK_EM:	
	333458	CH22_708FG_157_7_LINK_EM:	
	333611	CH22_872FG_217_6_LINK_EM:	
	333621	CH22_882FG_219_5_LINK_EM:	
	333814	CH22_1083FG_282_2_LINK_EM	
55	333849	CH22_1118FG_290_8_LINK_EM	
	335179	CH22_2515FG_504_9_LINK_EM	
	333949	CH22_1225FG_303_5_LINK_EM	
	333951	CH22_1227FG_303_7_LINK_EM	
	333955	CH22_1231FG_303_11_LINK_E	
60	335293	CH22_2635FG_527_6_LINK_EM	
	326816	c20_hs	
	326997	c21_hs	
	335550	CH22_2905FG_576_11_LINK_E	
	335581	CH22_2938FG_581_19_LINK_E	
65	335586	CH22_2944FG_581_25_LINK_E	

	328492	c_7_hs	
	335809	CH22_3181FG_617_6_LINK_EM	
	335810	CH22_3182FG_617_7_LINK_EM	
	335822	CH22_3195FG_619_7_LINK_EM	
5	335824	CH22_3197FG_619_11_LINK_E	
	335853	CH22_3228FG_626_5_LINK_EM	
	335886	CH22_3261FG_632_4_LINK_EM	
	330020	c16_p2	
	330211	c_5_p2	
10	337577	CH22_5864FG_LINK_C65E1.G	
	307848	AI364186	
	332797	CH22_13FG_6_2_LINK_C4G1.G	
	332798	CH22_14FG_6_5_LINK_C4G1.G	
	332799	CH22_15FG_6_6_LINK_C4G1.G	
15	334150	CH22_1429FG_339_1_LINK_EM	
	332933	CH22_154FG_38_7_LINK_C20H	
	332980	CH22_204FG_54_1_LINK_EM:A	
	332984	CH22_208FG_54_6_LINK_EM:A	
	334223	CH22_1507FG_360_4_LINK_EM	
20	334297	CH22_1588FG_372_3_LINK_EM	
	327098	c21_hs	
	334443	CH22_1742FG_387_2_LINK_EM	
	334444	CH22_1743FG_387_4_LINK_EM	
	334447	CH22_1746FG_387_7_LINK_EM	
25	334570	CH22_1875FG_405_11_LINK_E	
	334749	CH22_2061FG_427_1_LINK_EM	
	334777	CH22_2089FG_430_9_LINK_EM	
	336034	CH22_3419FG_678_5_LINK_DJ	
	334960	CH22_2281FG_465_29_LINK_E	
30	336441	CH22_3861FG_827_7_LINK_DJ	
	330551	9851_2	U39840 NM_004496 AW135607 BE087458 BE087567 AA177116 AW195705 AW750756 AI811008 AI694151 BE348594 AW971075 AI347950 AI201455 AI073898 AA652680 AA613671 AI318364 AA507550 AA693692 AI032599 AA991871 AI269801 AW948974 T74639 AA532907 AW949173
35	330786	53973_3	BE379594 AI192455 AL039862 AI744012 AI761735 AW243181 AI743687 AI928223 AI423022 AI627855 AI636059 AI651571 AW802044 AI826995 AI431733 AI539125 AA863056 AW270910 AI768930 AW008835 AW615183 AW591147 AI695294 AI672106 AA506358 AI308060 AA011556 AA962437 AI935488 BE219625 AI004356 AW151394 AI218466 N66178 AI419784 AW242519 AW946907 D60374 AA989263 AI698799 AA470460 AI824167
40	332247	372969_1	AA669097 AA513815 AA026798 AA676526 AA704429 AA704269 AW118292 AA579216 N58172
	332396	20265_1	AW579842 BE156562 BE156690 BE156489 BE081033 AK001559 BE149402 M85387 AW367811 AW367798 R17370 AI908947 AA382932 R58449 H18732 AA371231 AW962899 AA713530 AW892946 R53463 H11063 AW068542 Z40761 BE176212 BE176155 W23952 W92188 AW374883 AA303497 AW954769 AA036808 BE168063 AW382073 AW382085 AL041475 H80748 AI078161 BE463983 AI805213 AI761264 W94885 N94502 AI623772 AI419532 AI810302 AI634190 AW002516 AW150777 AI352312 AI367474 AW204807 AI675502 AI337026 AW134715 BE328451 AI123157 AI560020 AI300745 AI608631 AI248873 AA742484 AW051635 H18646 AI245045 AA507111 AI640510 AI925594 AA115747 AA143035 AA151106
45	332781	32044_1	AK001764 BE313896 AA380199 AA380151 AA194996 AW118089 AA495871 AW975219 AW085598 AI378909 AW992310 AW992409 AI911857 AA657643 AI804471 AI242589 AI623968 R09556 AI129100 AI206500 AA680094 AA677784 AI023178 AI277519 AA424742 AI240654 AA232846 AI804273 AI382376 AA001729 W90790 BE090656 AW295015 AI674596 AI431734 AI420517 AW769185 AI128355 AI192474 AI820001 AA001929 AA706925 AI076676 AI499119 AI200493 AI695919 AI376217 W69195 W69261 AW305099 W90320 BE048357 AI658856 AA838534 AA233258 AI753393 AA709227 AI674387 AI872616
50			

TABLE 3B shows the genomic positioning for those primekeys lacking unigene ID's and accession numbers in Table 3. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique number corresponding to an Eos probeset
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
 Strand: Indicates DNA strand from which exons were predicted.
 Nt_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	Nt_position
333611	Dunham, I. et al.	Plus	6548368-6548507
333621	Dunham, I. et al.	Plus	8597414-8597560
333814	Dunham, I. et al.	Plus	7894165-7894252
333849	Dunham, I. et al.	Plus	8018323-8018472
333949	Dunham, I. et al.	Plus	8589634-8589791
333951	Dunham, I. et al.	Plus	8592501-8592637
333955	Dunham, I. et al.	Plus	8597414-8597560
334150	Dunham, I. et al.	Plus	10529221-10529854
334297	Dunham, I. et al.	Plus	13420934-13421058
334443	Dunham, I. et al.	Plus	14298981-14299056
334444	Dunham, I. et al.	Plus	14306433-14306492
334447	Dunham, I. et al.	Plus	14308764-14308824
334570	Dunham, I. et al.	Plus	14994868-14994943
334777	Dunham, I. et al.	Plus	16259586-16260166
335179	Dunham, I. et al.	Plus	21634405-21634526
335581	Dunham, I. et al.	Plus	24976198-24976334
335586	Dunham, I. et al.	Plus	24990333-24990497
335809	Dunham, I. et al.	Plus	26310772-26310909
335810	Dunham, I. et al.	Plus	26314767-26314849
335822	Dunham, I. et al.	Plus	26364087-26364196
335824	Dunham, I. et al.	Plus	26376860-26376942
335886	Dunham, I. et al.	Plus	26934235-26934364
336034	Dunham, I. et al.	Plus	29014404-29014590
336441	Dunham, I. et al.	Plus	34187606-34187663
337577	Dunham, I. et al.	Plus	595377-595678
338260	Dunham, I. et al.	Plus	15458919-15459257
332797	Dunham, I. et al.	Minus	216964-216798
332798	Dunham, I. et al.	Minus	232147-231974
332799	Dunham, I. et al.	Minus	232421-232307
332933	Dunham, I. et al.	Minus	2035790-2035681
332980	Dunham, I. et al.	Minus	5136165-5136019
332984	Dunham, I. et al.	Minus	2632606-2632457
333168	Dunham, I. et al.	Minus	3729896-3729788
333169	Dunham, I. et al.	Minus	3730864-3730767
333452	Dunham, I. et al.	Minus	5136165-5136019
333456	Dunham, I. et al.	Minus	2631933-2631797
333458	Dunham, I. et al.	Minus	5143942-5143806
334223	Dunham, I. et al.	Minus	12734365-12734269
334749	Dunham, I. et al.	Minus	16090686-16090106
334960	Dunham, I. et al.	Minus	20160968-20160795
335293	Dunham, I. et al.	Minus	22316408-22316275
335550	Dunham, I. et al.	Minus	24668714-24668658
335853	Dunham, I. et al.	Minus	26614629-26614506
336624	Dunham, I. et al.	Minus	227714-227577
336625	Dunham, I. et al.	Minus	229124-229024
336679	Dunham, I. et al.	Minus	2035790-2035681
338255	Dunham, I. et al.	Minus	15242294-15242231
338561	Dunham, I. et al.	Minus	22311966-22311856
338562	Dunham, I. et al.	Minus	22312594-22312465
338759	Dunham, I. et al.	Minus	26582475-26582199
338763	Dunham, I. et al.	Minus	26628148-26628009
338764	Dunham, I. et al.	Minus	26641232-26641101

5	329960	5091594	Minus	1031-1162
	329929	6165201	Minus	156410-156553
10	330020	6671887	Plus	172397-172491
	326816	6552458	Plus	198354-198436
	326997	5867660	Minus	71389-72147
	327098	6682516	Minus	1061684-1062361
	330211	6013592	Plus	59158-59215
	328492	5868455	Minus	46094-46241
	329362	5868837	Minus	65688-68173

TABLE 4: shows a preferred subset of the Accession numbers for genes found in Table 3 which are differentially expressed in prostate tumor tissue compared to normal prostate tissue.

Pkey: Unique Eos probeset identifier number
 ExAccn: Exemplar Accession number, Genbank accession number
 UnigeneID: Unigene number
 Unigene Title: Unigene gene title
 R1: Ratio of tumor to normal body tissue

	Pkey	ExAccn	UnigeneID	Unigene Title	R1
15	100819	HG4020-HT4290	Hs.2387	Transglutaminase	10.5
	102698	U75272	Hs.1867	progastricsin (pepsinogen C)	10.6
	102869	X02544	Hs.572	orosomucoid 1	22.6
20	105370	AA236476	Hs.22791	ESTs; Weakly similar to transmembrane pr	10.3
	105645	AA282138	Hs.11325	ESTs	14
	106094	AA419461	Hs.23317	ESTs	10.9
	109014	AA156790	Hs.262036	ESTs	15.3
	109562	F01811	Hs.187931	ESTs; Moderately similar to voltage-gate	10.8
	113021	T23855	Hs.129836	KIAA1028 protein	10.8
25	114124	Z38595	Hs.125019	ESTs; Highly similar to KIAA0886 protein	21.3
	122791	AA460158	Hs.129836	KIAA1028 protein	12.4
	124352	N21626	Hs.102406	ESTs	10.2
	301042	AI659131	Hs.197733	ESTs	24.9
	302005	AI869666	Hs.123119	ESTs	36.8
30	302410	NM_004917	Hs.218366	EST cluster (not in UniGene) with exon h	26.8
	302881	AA508353	Hs.105314	relaxin 1 (H1)	78.8
	303344	AA255977	Hs.250646	ESTs; Highly similar to ubiquitin-conjug	19.5
	303753	AW503733	Hs.9414	ESTs	13
	310431	AI420227	Hs.149358	ESTs	72.9
35	311251	AI655662	Hs.197698	ESTs	41.3
	311596	AI682088	Hs.79375	ESTs	26.4
	312153	AA759250	Hs.118625	cytochrome b-561	11
	312521	AA033609	Hs.239884	ESTs	11.2
40	313676	AA861697	Hs.120591	EST cluster (not in UniGene)	13.4
	314171	AI821895	Hs.193481	ESTs	29.4
	314907	AI672225	Hs.222886	ESTs	19.3
	315051	AW292425	Hs.163484	EST	15.5
	315052	AA876910	Hs.134427	ESTs	20
	317548	AI654187	Hs.195704	ESTs	14.2
45	317869	AW295184	Hs.129142	ESTs; Weakly similar to DEOXYRIBONUCLEAS	13.8
	318428	AI949409	Hs.194591	ESTs	12.3
	318524	AW291511	Hs.159066	ESTs	25.9
	319080	Z45131	Hs.23023	ESTs	16.9
	319763	AA460775	Hs.6295	ESTs	14.3
50	320324	AF071202	Hs.139336	ATP-binding cassette; sub-family C (CFTR	56.2
	321441	AW297633	Hs.118498	ESTs	14.7
	322303	W07459	Hs.157601	EST cluster (not in UniGene)	22
	322782	AA056060	Hs.202577	EST cluster (not in UniGene)	18.4
	322818	AW043782	Hs.293616	ESTs	10.7
55	323287	AA639902	Hs.104215	ESTs	24.7
	324603	AW016378	Hs.292934	ESTs	24.2
	324617	AA508552	Hs.195839	ESTs	54
	324658	AI694767	Hs.129179	ESTs	22
	324691	AI217963	Hs.293341	ESTs; Weakly similar to Pro-a2(XI) [H.sa	10.6
60	324696	AA641092	Hs.257339	ESTs	10.2
	324718	AI557019	Hs.116467	ESTs	34.4
	330211		CH.05_p2 gil6013592		12.6
	330430	HG2261-HT2352	Hs.321110	Antigen, Prostate Specific, Alt. Splice	13.8
	330706	AA121140	Hs.177576	ESTs; Moderately similar to kynurenine a	14.5
65	330762	AA449677	Hs.15251	Human DNA sequence from clone 437M21 on	18.5
	330892	AA149579	Hs.91202	ESTs	15.3
	330949	H01458	Hs.142896	ESTs	10.3

5	331099	R36671	Hs.14846	ESTs	11.6
	331151	R82331	Hs.268838	ESTs	13
10	331889	AA431407	Hs.98802	Homo sapiens Chromosome 16 BAC clone CIT	33.6
	332247	N58172		ESTs	14.2
	332396	AA340504		ESTs; Weakly similar to similar to human	21.2
	332533	M99487	Hs.325825	folate hydrolase (prostate-specific memb	38.1
	332697	T94885	Hs.75725	carboxypeptidase E	24.3
	332797			CH22_FGENES.6_2	30.8
	332798			CH22_FGENES.6_5	66.8
	332799			CH22_FGENES.6_6	19.8
	334223			CH22_FGENES.360_4	20.3
	336624			CH22_FGENES.6-3	43.3
	336625			CH22_FGENES.6-4	37.9

TABLE 4A shows the accession numbers for those primekeys lacking unigeneID's for Table 4. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

10	Pkey:	Unique Eos probeset identifier number
	CAT number:	Gene cluster number
	Accession:	Genbank accession numbers
15	Pkey	CAT number
	336624	CH22_4071FG_6_3_
	336625	CH22_4072FG_6_4_
	330211	c_5_p2
20	332797	CH22_13FG_6_2_LINK_C4G1.G
	332798	CH22_14FG_6_5_LINK_C4G1.G
	332799	CH22_15FG_6_6_LINK_C4G1.G
	334223	CH22_1507FG_360_4_LINK_EM
25	332247	372969_1
	332396	20265_1
30	AA669097 AA513815 AA026798 AA676526 AA704429 AA704269 AW118292 AA579216 N58172 AW579842 BE156562 BE156690 BE156489 BE081033 AK001559 BE149402 M85387 AW367811 AW367798 R17370 AI908947 AA382932 R58449 H18732 AA371231 AW962899 AA713530 AW892946 R53463 H11063 AW068542 Z40761 BE176212 BE176155 W23952 W92188 AW374883 AA303497 AW954769 AA036808 BE168063 AW382073 AW382085 AL041475 H80748 AI078161 BE463983 AI805213 AI761264 W94885 N94502 AI623772 AI419532 AI810302 AI634190 AW002516 AW150777 AI352312 AI367474 AW204807 AI675502 AI337026 AW134715 BE328451 AI123157 AI560020 AI300745 AI608631 AI248873 AA742484 AW051635 H18646 AI245045 AA507111 AI640510 AI925594 AA115747 AA143035 AA151106	

TABLE 4B shows the genomic positioning for those primekeys lacking unigene ID's and accession numbers in Table 4. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique number corresponding to an Eos probeset
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
 Strand: Indicates DNA strand from which exons were predicted.
 Nt_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	Nt_position
332797	Dunham, I. et.al.	Minus	216964-216798
332798	Dunham, I. et.al.	Minus	232147-231974
332799	Dunham, I. et.al.	Minus	232421-232307
334223	Dunham, I. et.al.	Minus	12734365-12734269
336624	Dunham, I. et.al.	Minus	227714-227577
336625	Dunham, I. et.al.	Minus	229124-229024
330211	6013592	Plus	59158-59215

TABLE 5: 1170 GENES UP-REGULATED IN PROSTATE CANCER COMPARED TO NORMAL ADULT TISSUES

- 5 Table 5 shows 1170 genes up-regulated in prostate cancer compared to normal adult tissues. These were selected from 59680 probesets on the Affymetrix/Eos Hu03 GeneChip array such that the ratio of "average" prostate cancer to "average" normal adult tissues was greater than or equal to 3.44. The "average" prostate cancer level was set to the 85th percentile amongst 73 prostate cancers. The "average" normal adult tissue level was set to the 85th percentile amongst 162 non-malignant tissues. In order to remove gene-specific background levels of non-specific hybridization, the 7.5th percentile value amongst the 162 non-malignant tissues was subtracted from both the numerator and the denominator before the ratio was evaluated.

15	Pkey:	Unique Eos probeset identifier number			
	ExAccn:	Exemplar Accession number, Genbank accession number			
	UnigenelD:	Unigene number			
	Unigene Title:	Unigene gene title			
	R1:	Ratio of tumor to normal tissue			
20	Pkey	ExAccn	UnigenelD	Unigene Title	R1
	446057	AI420227	Hs.149358	ESTs, Weakly similar to A46010 X-linked	86.42
	400302	N48056	Hs.1915	folate hydrolase (prostate-specific memb	66.46
	414569	AF109298	Hs.118258	prostate cancer associated protein 1	58.36
25	417407	AA923278	Hs.290905	ESTs, Weakly similar to protease [H.sapi	56.16
	431579	AW971082	Hs.222886	ESTs, Weakly similar to TRHY_HUMAN TRICH	53.38
	409361	NM_005982	Hs.54416	sine oculis homeobox (Drosophila) homolo	48.28
	409731	AA125985	Hs.56145	thymosin, beta, identified in neuroblast	45.24
	400298	AA032279	Hs.61635	six transmembrane epithelial antigen of	43.48
30	420154	AI093155	Hs.95420	JM27 protein	41.12
	433466	AA508353	Hs.105314	relaxin 1 (H1)	39.88
	400296	AA305627	Hs.139336	ATP-binding cassette, sub-family C (CFTR	38.42
	400292	AA250737	Hs.72472	ESTs	38.00
	432887	AI926047	Hs.162859	ESTs	36.48
35	439176	AI446444	Hs.190394	ESTs, Weakly similar to B28096 line-1 pr	36.45
	430722	AW968543	Hs.203270	ESTs, Weakly similar to ALU1_HUMAN ALU S	33.20
	437052	AA861697	Hs.120591	ESTs	33.02
	418396	AI765805	Hs.26691	ESTs	32.68
	434036	AI659131	Hs.197733	hypothetical protein MGC2849	32.44
40	407709	AA456135	Hs.23023	ESTs	32.10
	426747	AA535210	Hs.171995	kallikrein 3, (prostate specific antigen	31.80
	407168	R45175		ESTs	31.72
	440260	AI972867	Hs.7130	copine IV	30.52
	421513	X00949	Hs.105314	relaxin 1 (H1)	30.10
45	416370	N90470	Hs.203697	ESTs, Weakly similar to I38022 hypotheti	29.68
	407122	H20276	Hs.31742	ESTs	29.24
	400287	S39329	Hs.181350	kallikrein 2, prostatic	28.90
	432244	AI669973	Hs.200574	ESTs	28.74
	451939	U80456	Hs.27311	single-minded (Drosophila) homolog 2	28.74
50	415989	AI267700	Hs.111128	ESTs	28.34
	418961	AW967646	Hs.23023	ESTs	27.34
	425628	NM_004476	Hs.1915	folate hydrolase (prostate-specific memb	27.32
	458509	AA654650	Hs.282906	ESTs	27.24
	448290	AK002107	Hs.20843	Homo sapiens cDNA FLJ11245 fis, clone PL	27.16
55	428336	AA503115	Hs.183752	microseminoprotein, beta-	26.17
	450096	AI682088	Hs.223368	holocarboxylase synthetase (biotin-[prop	25.60
	400299	X07730	Hs.171995	kallikrein 3, (prostate specific antigen	24.91
	437571	AA760894	Hs.153023	ESTs	24.74
	453160	AI263307	Hs.146228	H2B histone family, member L	24.66
60	453096	AW294631	Hs.11325	ESTs	24.46
	425075	AA506324	Hs.1852	acid phosphatase, prostate	24.23
	407202	N58172	Hs.109370	ESTs	24.18

	424846	AU077324	Hs.1832	neuropeptide Y	23.57
	453370	AI470523	Hs.182356	ATP-binding cassette, sub-family C (CFTR	23.16
	422805	AA436989	Hs.121017	H2A histone family, member A	22.52
5	444917	R68651	Hs.144997	ESTs	22.26
	408826	AF216077	Hs.48376	Homo sapiens clone HB-2 mRNA sequence	22.02
	413597	AW302885	Hs.117183	ESTs	21.76
	426429	X73114	Hs.169849	myosin-binding protein C, slow-type	21.32
	435981	H74319	Hs.188620	ESTs	21.12
10	432966	AA650114		ESTs	21.07
	418848	AI820961	Hs.193465	ESTs	21.06
	405685				20.90
	443271	BE568568	Hs.195704	ESTs	19.98
	418819	AA228776	Hs.191721	ESTs	19.94
15	420757	X78592	Hs.99915	androgen receptor (dihydrotestosterone r	19.72
	418994	AA296520	Hs.89546	selectin E (endothelial adhesion molecu	19.56
	429918	AW873986	Hs.119383	ESTs	19.04
	415539	AI733881	Hs.72472	ESTs	18.43
	450382	AA397658	Hs.60257	Homo sapiens cDNA FLJ13598 fis, clone PL	18.34
20	418829	AA516531	Hs.55999	NK homeobox (Drosophila), family 3, A	18.28
	429984	AL050102	Hs.227209	hypothetical protein FLJ21617	17.82
	443822	AI087412	Hs.143611	ESTs, Weakly similar to 2004399A chromos	17.66
	431676	AI685464	Hs.292638	gb:tt88f04.x1 NCI_CGAP_Pr28 Homo sapiens	17.64
	410330	AW023630	Hs.46786	ESTs	17.52
25	432441	AW292425	Hs.163484	ESTs	17.41
	452792	AB037765	Hs.30652	KIAA1344 protein	17.39
	445472	AB006631	Hs.12784	Homo sapiens mRNA for KIAA0293 gene, par	17.00
	414565	AA502972	Hs.183390	hypothetical protein FLJ13590	16.82
	430487	D87742	Hs.241552	KIAA0268 protein	16.72
30	431716	D89053	Hs.268012	fatty-acid-Coenzyme A ligase, long-chain	16.60
	419536	AA603305		gb:np12d11.s1 NCI_CGAP_Pr3 Homo sapiens	16.50
	439677	R82331	Hs.164599	ESTs	16.46
	449625	NM_014253	Hs.23796	odz (odd Oz/ten-m, Drosophila) homolog 1	16.32
	408430	S79876	Hs.44926	dipeptidylpeptidase IV (CD26, adenosine	16.28
35	447033	AI357412	Hs.157601	ESTs	16.02
	453006	AI362575	Hs.167133	ESTs	15.74
	431474	AL133990	Hs.190642	ESTs	15.70
	420218	AW958037	Hs.22437	ribosomal protein L4	15.64
	408000	L11690	Hs.620	bullous pemphigoid antigen 1 (230/240kD)	15.54
40	416208	AW291168	Hs.41295	ESTs, Weakly similar to MUC2_HUMAN MUCIN	15.48
	430226	BE245562	Hs.2551	adrenergic, beta-2-, receptor, surface	15.40
	415263	AA948033	Hs.130853	ESTs	15.38
	432437	W07088	Hs.293685	ESTs	15.26
	428398	AI249368	Hs.98558	ESTs	15.21
45	429900	AA460421	Hs.30875	ESTs	14.90
	449156	AF103907	Hs.171353	prostate cancer antigen 3	14.89
	411096	U80034	Hs.68583	mitochondrial intermediate peptidase	14.81
	435974	U29690	Hs.37744	Homo sapiens beta-1 adrenergic receptor	14.76
	444484	AK002126	Hs.11260	hypothetical protein FLJ11264	14.76
50	422728	AW937826	Hs.103262	ESTs, Weakly similar to ZN91_HUMAN ZINC	14.60
	418601	AA279490	Hs.86368	calmegin	14.56
	448999	AF179274	Hs.22791	transmembrane protein with EGF-like and	14.55
	445885	AI734009	Hs.127699	KIAA1603 protein	14.44
	452712	AW838616		gb:RC5-LT0054-140200-013-D01 LT0054 Homo	14.22
55	432189	AA527941		gb:nh30c04.s1 NCI_CGAP_Pr3 Homo sapiens	14.12
	424565	AW102723	Hs.75295	guanylate cyclase 1, soluble, alpha 3	13.78
	429290	AF203032	Hs.198760	neurofilament, heavy polypeptide (200kD)	13.57
	419264	AA877104	Hs.293672	ESTs, Weakly similar to ALUB_HUMAN !!!!	13.40
	416445	AL043004	Hs.300678	KIAA0135 protein	13.32
60	407275	AI364186		gb:qw34h07.x1 NCI_CGAP_UT4 Homo sapiens	13.24
	408369	R38438	Hs.182575	solute carrier family 15 (H+/peptide tra	13.21
	446720	AI439136	Hs.140546	ESTs	13.06
	434988	AI418055	Hs.161160	ESTs	13.02
	448172	N75276	Hs.135904	ESTs	12.98
65	416182	NM_004354	Hs.79069	cyclin G2	12.94
	420544	AA677577	Hs.98732	Homo sapiens Chromosome 16 BAC clone CIT	12.79
	445413	AA151342	Hs.12677	CGI-147 protein	12.64
	452588	AA889120	Hs.110637	homeo box A10	12.62
	407819	R42185	Hs.274803	ESTs	12.60
	433444	AW975324	Hs.129816	ESTs	12.60

	421059	AI654133	Hs.30212	thyroid receptor interacting protein 15	12.30
	420077	AW512260	Hs.87767	ESTs	12.24
	453930	AA419466	Hs.36727	hypothetical protein FLJ10903	12.22
5	441610	AW576148	Hs.148376	ESTs	12.20
	451009	AA013140	Hs.115707	ESTs	12.18
	433764	AW753676	Hs.39982	ESTs	12.16
	440286	U29589	Hs.7138	cholinergic receptor, muscarinic 3	12.04
	443912	R37257	Hs.184780	ESTs	11.92
	419526	AI821895	Hs.193481	ESTs	11.91
10	423073	BE252922	Hs.123119	MAD (mothers against decapentaplegic, Dr	11.87
	452784	BE463857	Hs.151258	hypothetical protein FLJ21062	11.86
	414422	AA147224	Hs.71814	ESTs	11.76
	450203	AF097994	Hs.301528	L-kynurenine/alpha-aminoadipate aminotra	11.68
	436679	AI127483	Hs.120451	ESTs, Weakly similar to unnamed protein	11.60
15	440901	AA909358	Hs.128612	ESTs	11.60
	448045	AJ297436	Hs.20166	prostate stem cell antigen	11.51
	433887	AW204232	Hs.279522	ESTs	11.50
	434980	AW770553	Hs.293640	sterol O-acyltransferase (acyl-Coenzyme	11.38
20	425905	AB032959	Hs.161700	novel C3HC4 type Zinc finger (ring finge	11.33
	434680	T11738	Hs.127574	ESTs	11.32
	449650	AF055575	Hs.297647	calcium channel, voltage-dependent, L ty	11.18
	431173	AW971198	Hs.294068	ESTs	11.16
	434539	AW748078	Hs.214410	ESTs, Weakly similar to MUC2_HUMAN MUCIN	11.16
	410037	AB020725	Hs.58009	KIAA0918 protein	11.14
25	417708	N74392	Hs.50495	ESTs	11.14
	458332	AI000341	Hs.220491	ESTs	11.12
	420381	D50640	Hs.301782	phosphodiesterase 3B, cGMP-inhibited	11.10
	425665	AK001050	Hs.159066	hypothetical protein FLJ10188	11.08
	425710	AF030880	Hs.159275	solute carrier family, member 4	11.08
30	428728	NM_016625	Hs.191381	hypothetical protein	11.04
	407021	U52077		gb:Human mariner1 transposase gene, comp	11.02
	410733	D84284	Hs.66052	CD38 antigen (p45)	11.02
	401714				10.90
	434485	AI623511	Hs.118567	ESTs	10.89
35	415786	AW419196	Hs.257924	hypothetical protein FLJ13782	10.87
	452340	NM_002202	Hs.505	ISL1 transcription factor, LIM/homeodoma	10.85
	453628	AW243307	Hs.170187	hypothetical protein	10.72
	408063	BE086548	Hs.42346	calcineurin-binding protein calsarcin-1	10.67
	417687	AI828596	Hs.250691	ESTs	10.64
40	434666	AF151103	Hs.112259	T cell receptor gamma locus	10.53
	432374	W68815	Hs.301885	Homo sapiens cDNA FLJ11346 fis, clone PL	10.50
	428819	AL135623	Hs.193914	KIAA0575 gene product	10.48
	413409	AI638418	Hs.21745	DEAD/H (Asp-Glu-Ala-Asp/His) box polypep	10.44
	428775	AA434579	Hs.143691	ESTs	10.21
45	436556	AI364997	Hs.7572	ESTs	10.20
	441690	R81733	Hs.33106	ESTs	10.14
	419852	AW503756	Hs.286184	hypothetical protein dJ551D2.5	10.10
	421991	NM_014918	Hs.110488	KIAA0990 protein	10.04
	423698	AA329796	Hs.1098	DKFZp434J1813 protein	10.02
50	452039	AI922988	Hs.172510	ESTs	10.00
	433043	W57554	Hs.125019	ESTs	9.98
	433927	AI557019	Hs.116467	small nuclear protein PRAC	9.97
	445424	AB028945	Hs.12696	cortactin SH3 domain-binding protein	9.96
	432240	AI694767	Hs.129179	Homo sapiens cDNA FLJ13581 fis, clone PL	9.88
55	433104	AL043002	Hs.128246	ESTs, Moderately similar to unnamed prot	9.84
	452744	AI267652	Hs.30504	Homo sapiens mRNA; cDNA DKFZp434E082 (fr	9.82
	431217	NM_013427	Hs.250830	Rho GTPase activating protein 6	9.75
	427398	AW390020	Hs.20415	chromosome 21 open reading frame 11	9.70
	446896	T15767	Hs.22452	Homo sapiens mRNA for KIAA1737 protein,	9.70
60	421470	R27496	Hs.1378	annexin A3	9.64
	406554				9.60
	401424				9.58
	407902	AL117474	Hs.41181	Homo sapiens mRNA; cDNA DKFZp727C191 (fr	9.56
	423545	AP000692	Hs.129781	chromosome 21 open reading frame 5	9.54
65	439024	R96696	Hs.35598	ESTs	9.51
	431548	AI834273	Hs.9711	novel protein	9.48
	409262	AK000631	Hs.52256	hypothetical protein FLJ20624	9.45
	446271	D82484	Hs.100469	ESTs	9.42
	448692	AW013907	Hs.224276	methylcrotonoyl-Coenzyme A carboxylase 2	9.26

	414140	AA281279	Hs.23317	hypothetical protein FLJ14681	9.24
	435980	AF274571	Hs.129142	deoxyribonuclease II beta	9.24
	421246	AW582962	Hs.300961	CGI-47 protein	9.20
	427304	AA761526	Hs.163853	ESTs	9.16
5	442914	AW188551	Hs.99519	hypothetical protein FLJ14007	9.16
	413627	BE182082	Hs.246973	ESTs	9.14
	439699	AF086534	Hs.187561	ESTs, Moderately similar to ALU1_HUMAN A	9.10
	437718	AI927288	Hs.196779	ESTs	9.07
	439820	AL360204	Hs.283853	Homo sapiens mRNA full length insert cDN	9.06
10	447342	AI199268	Hs.19322	Homo sapiens, Similar to RIKEN cDNA 2010	9.05
	446223	BE300091	Hs.119699	hypothetical protein FLJ12969	9.04
	410001	AB041036	Hs.57771	kallikrein 11	9.03
	424012	AW368377	Hs.137569	tumor protein 63 kDa with strong homolog	9.03
	441791	AW372449	Hs.175982	hypothetical protein FLJ21159	9.02
15	448206	BE622585	Hs.3731	ESTs, Moderately similar to I38022 hypot	9.02
	414269	AA298489		olfactory receptor, family 51, subfamily	8.99
	442081	AA401863	Hs.22380	ESTs	8.98
	420092	AA814043	Hs.88045	ESTs	8.85
	411630	U42349	Hs.71119	Putative prostate cancer tumor suppresso	8.80
20	421863	AI952677	Hs.108972	Homo sapiens mRNA; cDNA DKFZp434P228 (fr	8.80
	454141	AW138413	Hs.182356	ATP-binding cassette, sub-family C (CFTR	8.80
	418278	AI088489	Hs.83937	hypothetical protein	8.78
	428330	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin,	8.76
	432415	T16971	Hs.289014	ESTs, Weakly similar to A43932 mucin 2 p	8.75
25	424906	AI566086	Hs.153716	Homo sapiens mRNA for Hmob33 protein, 3'	8.74
	415245	N59650	Hs.27252	ESTs	8.72
	442409	BE208843	Hs.129544	hypothetical protein MGC15438	8.70
	404571				8.66
	418033	W68180	Hs.259855	elongation factor-2 kinase	8.64
30	456497	AW967956	Hs.123648	ESTs, Weakly similar to AF108460 1 ubinu	8.56
	405876				8.54
	448807	AI571940	Hs.7549	ESTs	8.52
	445372	N36417	Hs.144928	ESTs	8.48
	425171	AW732240	Hs.300615	ESTs	8.44
35	419968	X04430	Hs.93913	interleukin 6 (interferon, beta 2)	8.36
	407385	AA610150	Hs.272072	ESTs, Weakly similar to I38022 hypotheti	8.31
	433172	AB037841	Hs.102652	hypothetical protein ASH1	8.30
	422631	BE218919	Hs.118793	hypothetical protein FLJ10688	8.27
	412719	AW016610	Hs.129911	ESTs	8.24
40	418849	AW474547	Hs.53565	Homo sapiens PIG-M mRNA for mannosyltran	8.22
	444922	AI921750	Hs.144871	Homo sapiens cDNA FLJ13752 fis, clone PL	8.22
	427674	NM_003528	Hs.2178	H2B histone family, member Q	8.20
	432101	AI918950	Hs.11092	EphA3	8.17
	416288	H51299		gb:yp07c06.s1 Soares breast 3NbHBst Homo	8.15
45	404915				8.08
	440106	AA864968	Hs.127699	KIAA1603 protein	8.07
	442861	AA243837	Hs.57787	ESTs	8.06
	452259	AA317439	Hs.28707	signal sequence receptor, gamma (translo	8.06
	443250	AI041530	Hs.132107	ESTs	8.06
50	437267	AW511443	Hs.258110	ESTs	8.04
	452891	N75582	Hs.212875	ESTs, Weakly similar to DYH9_HUMAN CILI	8.02
	422219	AW978073		regulator of mitotic spindle assembly 1	8.00
	453049	BE537217	Hs.30343	ESTs	8.00
	439731	AI953135	Hs.45140	hypothetical protein FLJ14084	7.98
55	408554	AA836381	Hs.7323	nuclear receptor co-repressor/HDAC3 comp	7.94
	421154	AA284333	Hs.287631	Homo sapiens cDNA FLJ14269 fis, clone PL	7.94
	430107	AA465293	Hs.105069	ESTs	7.94
	433404	T32982	Hs.102720	ESTs	7.93
	450813	AI739625	Hs.203376	ESTs	7.90
60	416239	AL038450	Hs.48948	ESTs	7.85
	448212	AI475858		gb:tc87d07.x1 NCI_CGAP_CLL1 Homo sapiens	7.82
	449532	W74653	Hs.271593	ESTs, Moderately similar to A47582 B-cel	7.82
	413930	M86153	Hs.75618	RAB11A, member RAS oncogene family	7.80
	458191	AI420611	Hs.127832	ESTs	7.80
65	444858	AI199738	Hs.208275	ESTs, Weakly similar to ALUA_HUMAN !!!!	7.78
	457498	AI732230	Hs.191737	ESTs	7.78
	407235	D20569	Hs.169407	SAC2 (suppressor of actin mutations 2, y	7.76
	433759	AA680003	Hs.109363	Homo sapiens cDNA: FLJ23603 fis, clone L	7.74
	433805	AA706910	Hs.112742	ESTs	7.74

	426485	NM_006207	Hs.170040	platelet-derived growth factor receptor-	7.72
	446028	R44714	Hs.106795	Homo sapiens cDNA FLJ13136 fis, clone NT	7.72
	418555	AI417215	Hs.87159	hypothetical protein FLJ12577	7.70
	447499	AW262580	Hs.147674	protocadherin beta 16	7.70
5	419839	U24577	Hs.93304	phospholipase A2, group VII (platelet-ac	7.68
	416857	AA188775	Hs.292453	ESTs	7.68
	413801	M62246	Hs.35406	ESTs, Highly similar to unnamed protein	7.66
	425480	AB023198	Hs.158135	KIAA0981 protein	7.66
	420120	AL049610	Hs.95243	transcription elongation factor A (SII)-	7.64
10	424099	AF071202	Hs.139336	ATP-binding cassette, sub-family C (CFTR	7.64
	446307	T50083	Hs.9094	ESTs	7.63
	429220	AW207206	Hs.136319	ESTs	7.59
	420345	AW295230	Hs.25231	ESTs	7.54
	429208	AA447990	Hs.190478	ESTs	7.54
15	447247	AW369351	Hs.287955	Homo sapiens cDNA FLJ13090 fis, clone NT	7.53
	440995	T57773	Hs.10263	ESTs	7.53
	448706	AW291095	Hs.21814	interleukin 20 receptor, alpha	7.52
	410227	AB009284	Hs.61152	exostoses (multiple)-like 2	7.49
	431616	AA508552	Hs.195839	ESTs, Weakly similar to I38022 hypotheti	7.46
20	434217	AW014795	Hs.23349	ESTs	7.44
	431467	N71831	Hs.256398	Homo sapiens mRNA; cDNA DKFZp434E0528 (f	7.42
	448519	AW175665	Hs.244334	Homo sapiens prostatein mRNA, complete cds	7.42
	446791	AI632278	Hs.34981	ESTs	7.40
	419743	AW408762	Hs.127478	Homo sapiens clone 24416 mRNA sequence	7.39
25	445855	BE247129	Hs.145569	ESTs	7.36
	425211	M18667	Hs.1867	progastricsin (pepsinogen C)	7.35
	419131	AA406293	Hs.301622	ESTs	7.34
	400294	N95796	Hs.179809	Homo sapiens prostatein mRNA, complete cds	7.33
	441736	AW292779	Hs.169799	ESTs	7.28
30	427701	AA411101	Hs.221750	nuclear autoantigenic sperm protein (his	7.24
	457733	AW974812	Hs.291971	ESTs	7.24
	418432	M14156	Hs.85112	insulin-like growth factor 1 (somatomedi	7.22
	441201	AW118822	Hs.128757	ESTs	7.21
	419953	BE267154	Hs.125752	ESTs	7.20
35	419991	AJ000098	Hs.94210	eyes absent (Drosophila) homolog 1	7.20
	425018	BE245277	Hs.154196	E4F transcription factor 1	7.20
	424560	AA158727	Hs.150555	protein predicted by clone 23733	7.18
	435380	AA679001	Hs.192221	ESTs	7.14
	420658	AW965215	Hs.130707	ESTs	7.12
40	408291	AB023191	Hs.44131	KIAA0974 protein	7.10
	409110	AA191493	Hs.48778	niban protein	7.10
	414485	W27026	Hs.182625	VAMP (vesicle-associated membrane protei	7.10
	430039	BE253012	Hs.153400	ESTs, Weakly similar to ALU1_HUMAN ALU S	7.10
	450832	AW970602	Hs.105421	ESTs	7.10
45	417153	X57010	Hs.81343	collagen, type II, alpha 1 (primary oste	7.08
	412446	AI768015	Hs.92127	ESTs	7.07
	412953	Z45794	Hs.238809	ESTs	7.06
	418051	AW192535	Hs.19479	ESTs	7.06
	421566	NM_000399	Hs.1395	early growth response 2 (Krox-20 (Drosop	7.04
50	446999	AA151520	Hs.279525	hypothetical protein MGC4485	7.04
	440529	AW207640	Hs.16478	Homo sapiens cDNA: FLJ21718 fis, clone C	7.04
	441111	AI806867	Hs.126594	ESTs	7.01
	451027	AW519204	Hs.40808	ESTs	7.00
	408432	AW195262		gb:xn67b05.x1 NCL_CGAP_CML1 Homo sapiens	7.00
55	432223	AA333283	Hs.285336	Homo sapiens, clone IMAGE:3460280, mRNA	7.00
	444805	AB007899	Hs.12017	homolog of yeast ubiquitin-protein ligas	6.99
	414212	AA136569	Hs.295940	KIAA0187 gene product	6.98
	431725	X65724	Hs.2839	Norrie disease (pseudoglioma)	6.98
	449685	AW296669	Hs.66095	ESTs	6.97
60	447313	U92981	Hs.18081	Homo sapiens clone DT1P1B6 mRNA, CAG rep	6.96
	424590	AW966399	Hs.46821	hypothetical protein FLJ20086	6.94
	449655	AI021987	Hs.59970	ESTs	6.92
	419563	AA526235	Hs.193162	Homo sapiens cDNA FLJ11983 fis, clone HE	6.90
	434163	AW974720	Hs.25206	group XII secreted phospholipase A2	6.89
65	415809	Z32789	Hs.46601	ESTs	6.86
	425782	U66468	Hs.159525	cell growth regulatory with EF-hand doma	6.85
	417958	AA767382	Hs.193417	ESTs	6.84
	427408	AA583206	Hs.2156	RAR-related orphan receptor A	6.79
	445873	AA250970	Hs.251946	poly(A)-binding protein, cytoplasmic 1-I	6.74

	410718	AI920783	Hs.191435	ESTs	6.74
	432363	AA534489		gb:nf76g11.s1 NCI_CGAP_Co3 Homo sapiens	6.74
	436521	AW203986	Hs.213003	ESTs	6.73
5	435604	AA625279	Hs.26892	uncharacterized bone marrow protein BM04	6.73
	419083	AI479560	Hs.98613	Homo sapiens cDNA FLJ12292 fis, clone MA	6.72
	418245	AA088767	Hs.83883	transmembrane, prostate androgen induced	6.70
	420714	BE172704	Hs.222746	KIAA1610 protein	6.70
	412707	AW206373	Hs.16443	Homo sapiens cDNA: FLJ21721 fis, clone C	6.67
10	421896	N62293	Hs.45107	ESTs	6.66
	411078	AI222020	Hs.182364	CocoaCrisp	6.66
	452465	AA610211	Hs.34244	ESTs	6.66
	422763	AA033699	Hs.83938	ESTs, Moderately similar to MAS2_HUMAN M	6.66
	444618	AV653785	Hs.300171	ELL-RELATED RNA POLYMERASE II, ELONGATIO	6.64
15	450164	AI239923	Hs.30098	ESTs	6.63
	431060	AF039307	Hs.249171	homeo box A11	6.62
	408031	AA081395	Hs.42173	Homo sapiens cDNA FLJ10366 fis, clone NT	6.62
	420285	AA258124	Hs.293878	ESTs, Moderately similar to ZN91_HUMAN Z	6.62
	444670	H58373	Hs.37494	hypothetical protein MGC5370	6.62
20	444489	AI151010	Hs.157774	ESTs	6.60
	445685	AW779829	Hs.263436	gb:hn88a05.x1 NCI_CGAP_Kid11 Homo sapien	6.60
	435677	AA694142	Hs.293726	ESTs, Weakly similar to TSGA RAT TESTIS	6.59
	452221	C21322	Hs.11577	hypothetical protein FLJ22242	6.59
	431510	AA580082	Hs.112264	ESTs	6.56
25	415874	AF091622	Hs.78893	KIAA0244 protein	6.54
	418405	AI868282	Hs.11898	ESTs, Highly similar to KIAA1370 protein	6.54
	452768	AW069459	Hs.61539	ESTs	6.54
	401451				6.52
	416289	W26333		ESTs	6.52
30	431778	AL080276	Hs.268562	regulator of G-protein signalling 17	6.51
	409089	NM_014781	Hs.50421	KIAA0203 gene product	6.50
	442833	AA328153	Hs.88201	ESTs, Weakly similar to A Chain A, Cryst	6.50
	431992	NM_002742	Hs.2891	protein kinase C, mu	6.49
	418833	AW974899	Hs.292776	ESTs	6.48
35	429163	AA884766		gb:am20a10.s1 Soares_NFL_T_GBC_S1 Homo s	6.46
	430403	AF039390	Hs.241382	tumor necrosis factor (ligand) superfami	6.46
	443058	AW451642	Hs.16732	ESTs	6.46
	418564	AA631143	Hs.179809	Homo sapiens prostein mRNA, complete cds	6.44
	432674	AA641092	Hs.257339	ESTs, Weakly similar to I38022 hypotheti	6.44
40	423600	AI633559	Hs.29076	ESTs	6.44
	404253				6.42
	433610	AA806822	Hs.112547	ESTs	6.42
	421552	AF026692	Hs.105700	secreted frizzled-related protein 4	6.41
	407118	AA156790	Hs.262036	ESTs, Weakly similar to Z223_HUMAN ZINC	6.40
45	408608	N79738	Hs.136102	KIAA0853 protein	6.40
	421452	AI925946	Hs.104530	fetal hypothetical protein	6.40
	433285	AW975944	Hs.237396	ESTs	6.40
	434926	BE543269	Hs.50252	mitochondrial ribosomal protein L32	6.40
	446189	H85224	Hs.214013	ESTs	6.40
50	416806	NM_000288	Hs.79993	peroxisomal biogenesis factor 7	6.38
	416467	H57585	Hs.37467	ESTs	6.36
	453403	BE466639	Hs.61779	Homo sapiens cDNA FLJ13591 fis, clone PL	6.34
	429769	NM_004917	Hs.218366	kallikrein 4 (protease, enamel matrix, p	6.34
	423642	AW452650	Hs.157148	hypothetical protein MGC13204	6.32
55	425843	BE313280	Hs.159627	death associated protein 3	6.32
	439221	AA737106	Hs.32250	ESTs, Moderately similar to I78885 serin	6.32
	428194	AA765603	Hs.180877	H3 histone, family 3B (H3.3B)	6.30
	431958	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta	6.30
	439366	AF100143	Hs.6540	fibroblast growth factor 13	6.30
60	452789	AW081626	Hs.242561	ESTs	6.30
	416836	D54745	Hs.80247	cholecystokinin	6.30
	436962	AW377314	Hs.5364	DKFZP564I052 protein	6.29
	433383	AF034837	Hs.192731	double-stranded RNA specific adenosine d	6.29
	418636	AW749855		gb:QV4-BT0534-281299-053-c05 BT0534 Homo	6.26
65	450728	AW162923	Hs.25363	presenilin 2 (Alzheimer disease 4)	6.25
	440293	AI004193	Hs.22123	ESTs	6.24
	453745	AA952989	Hs.63908	hypothetical protein MGC14726	6.24
	426595	AW971980	Hs.62402	p21/Cdc42/Rac1-activated kinase 1 (yeast	6.24
	444412	AI147652	Hs.216381	Homo sapiens clone HH409 unknown mRNA	6.24
	413384	NM_000401	Hs.75334	exostoses (multiple) 2	6.22

	426320	W47595	Hs.169300	transforming growth factor, beta 2	6.22
	423349	AF010258	Hs.127428	homeo box A9	6.20
	429165	AW009886	Hs.118258	prostate cancer associated protein 1	6.18
	424800	AL035588	Hs.153203	MyoD family inhibitor	6.18
5	409564	AA045857	Hs.54943	fracture callus 1 (rat) homolog	6.16
	438796	W67821	Hs.109590	genethonin 1	6.16
	425451	AF242769	Hs.157461	mesenchymal stem cell protein DSC54	6.14
	451663	AI872360	Hs.209293	ESTs	6.14
10	413623	AA825721	Hs.246973	ESTs	6.12
	452232	AW020603	Hs.271698	radial spoke protein 3	6.12
	453390	AA862496	Hs.28482	ESTs	6.12
	435542	AA687376	Hs.269533	ESTs	6.12
	420424	AB033036	Hs.97594	KIAA1210 protein	6.11
	407103	AA424881	Hs.256301	hypothetical protein MGC13170	6.10
15	409734	BE161664	Hs.56155	hypothetical protein	6.10
	432686	BE223007	Hs.152460	Homo sapiens cDNA FLJ12909 fis, clone NT	6.10
	438361	AA805666	Hs.146217	Homo sapiens cDNA: FLJ23077 fis, clone L	6.10
	411479	AW848047		gb:IL3-CT0214-291299-052-A12 CT0214 Homo	6.10
	438849	W28948	Hs.10762	ESTs	6.08
20	452726	AF188527	Hs.61661	ESTs, Weakly similar to AF174605 1 F-box	6.08
	445895	D29954	Hs.13421	KIAA0056 protein	6.08
	440774	AI420611	Hs.127832	ESTs	6.07
	422583	AA410506	Hs.118578	KIAA0874 protein	6.06
25	427500	AW970017	Hs.293948	ESTs, Weakly similar to S65657 alpha-1C-	6.04
	443646	AI085198	Hs.298699	ESTs	6.04
	410566	AA373210	Hs.43047	Homo sapiens cDNA FLJ13585 fis, clone PL	6.02
	417845	AL117461	Hs.82719	Homo sapiens mRNA; cDNA DKFZp586F1822 (f	6.02
	430273	AI311127	Hs.125522	ESTs	6.02
	434792	AA649253	Hs.132458	ESTs	6.01
30	442490	AW965078	Hs.30212	thyroid receptor interacting protein 15	6.01
	420026	AI831190	Hs.166676	ESTs	6.00
	437782	AI370876	Hs.123163	exportin 1 (CRM1, yeast, homolog)	6.00
	447359	NM_012093	Hs.18268	adenylate kinase 5	6.00
	447713	AI420733	Hs.207083	ESTs	6.00
35	451073	AI758905	Hs.206063	ESTs	6.00
	451640	AA195601	Hs.26771	Human DNA sequence from clone 747H23 on	6.00
	410889	X91662	Hs.66744	twist (Drosophila) homolog (acrocephalos	5.97
	441222	AI277237	Hs.44208	hypothetical protein FLJ23153	5.96
40	447732	AI758398	Hs.161318	ESTs	5.96
	437756	AA767537	Hs.197096	ESTs	5.95
	408829	NM_006042	Hs.48384	heparan sulfate (glucosamine) 3-O-sulfot	5.94
	453911	AW503857	Hs.4007	Sarcolemmal-associated protein	5.94
	414085	AA114016	Hs.75746	aldehyde dehydrogenase 1 family, member	5.93
45	408875	NM_015434	Hs.48604	DKFZP434B168 protein	5.92
	439451	AF086270	Hs.278554	heterochromatin-like protein 1	5.92
	423853	AB011537	Hs.133466	slit (Drosophila) homolog 1	5.91
	453060	AW294092	Hs.21594	hypothetical protein MGC15754	5.91
	420407	AA814732	Hs.145010	lipopolysaccharide-specific response 5-ii	5.91
	450480	X82125	Hs.25040	zinc finger protein 239	5.90
50	408446	AW450669	Hs.45068	hypothetical protein DKFZp434I143	5.88
	421039	NM_003478	Hs.101299	cullin 5	5.88
	451684	AF216751	Hs.26813	CDA14	5.88
	436063	AK000028	Hs.250867	ribosomal protein S24	5.86
	410507	AA355288	Hs.271408	transitional epithelia response protein	5.86
55	420179	N74530	Hs.21168	ESTs	5.84
	453878	AW964440	Hs.19025	DC32	5.84
	452270	AW975014	Hs.26	ferrochelatase (protoporphyrin)	5.83
	435867	AA954229	Hs.114052	ESTs	5.82
	417683	AW566008	Hs.239154	ankyrin repeat, family A (RFXANK-like),	5.82
60	432005	AA524190	Hs.120777	ESTs, Weakly similar to ELL2_HUMAN RNA P	5.81
	406815	AA833930	Hs.288036	tRNA isopentenylpyrophosphate transferas	5.80
	437980	R50393	Hs.278436	KIAA1474 protein	5.80
	425856	AA364908	Hs.98927	hypothetical protein FLJ13993	5.79
	400301	X03635	Hs.1657	estrogen receptor 1	5.78
65	446261	AA313893	Hs.13399	hypothetical protein FLJ12615 similar to	5.78
	410141	R07775	Hs.287657	Homo sapiens cDNA: FLJ21291 fis, clone C	5.77
	427258	AA400091	Hs.39421	ESTs	5.76
	419108	AA389724	Hs.191264	ESTs, Weakly similar to ALU7_HUMAN ALU S	5.76
	442029	AW956698	Hs.14456	neural precursor cell expressed, develop	5.76

	407783	AW996872	Hs.172028	a disintegrin and metalloproteinase doma	5.75
	434408	AI031771	Hs.132586	ESTs	5.74
	415077	L41607	Hs.934	glucosaminyl (N-acetyl) transferase 2, I	5.74
	432435	BE218886	Hs.282070	ESTs	5.74
5	433313	W20128	Hs.296039	ESTs	5.73
	431740	N75450	Hs.183412	ESTs, Moderately similar to AF116721 67	5.73
	412991	AW949013		gb:QV4-FT0005-110500-201-e12 FT0005 Homo	5.72
	418852	BE537037	Hs.273294	hypothetical protein FLJ20069	5.72
10	418882	NM_004996	Hs.89433	ATP-binding cassette, sub-family C (CFTR	5.72
	446867	AB007891	Hs.16349	KIAA0431 protein	5.72
	437866	AA156781	Hs.83992	metallothionein 1E (functional)	5.72
	410232	AW372451	Hs.61184	CGI-79 protein	5.70
	414452	AA454038	Hs.29032	ESTs	5.70
15	422762	AL031320	Hs.119976	Human DNA sequence from clone RP1-20N2 o	5.70
	428730	AA625947	Hs.25750	ESTs	5.70
	431571	AW500486	Hs.180610	splicing factor proline/glutamine rich (5.70
	433393	AF038564	Hs.98074	itchy (mouse homolog) E3 ubiquitin prote	5.70
	450616	AL133067	Hs.25214	hypothetical protein	5.70
20	443774	AL117428	Hs.9740	DKFZP434A236 protein	5.69
	446100	AW967109	Hs.13804	hypothetical protein dJ462O23.2	5.69
	419168	AI336132	Hs.33718	Homo sapiens cDNA FLJ12641 fis, clone NT	5.68
	416653	AA768553	Hs.77496	metallothionein 1E (functional)	5.67
	452679	Z42387	Hs.4299	transmembrane, prostate androgen induced	5.66
25	450244	AA007534	Hs.125062	ESTs	5.66
	408621	AI970672	Hs.46638	chromosome 11 open reading frame 8	5.65
	450325	AI935962	Hs.26289	ESTs	5.65
	439671	AW162840	Hs.6641	kinesin family member 5C	5.64
	452387	AI680772	Hs.4316	trinucleotide repeat containing 12	5.64
30	413992	W26276	Hs.136075	RNA, U2 small nuclear	5.63
	444151	AW972917	Hs.128749	alpha-methylacyl-CoA racemase	5.63
	417791	AW965339	Hs.111471	ESTs	5.62
	410196	AI936442	Hs.59838	hypothetical protein FLJ10808	5.60
	415123	D60925		ESTs	5.60
35	429170	NM_001394	Hs.2359	dual specificity phosphatase 4	5.60
	434415	BE177494		gb:RC6-HT0596-270300-011-C05 HT0596 Homo	5.60
	440738	AI004650	Hs.225674	WD repeat domain 9	5.60
	443830	AI142095	Hs.143273	ESTs	5.60
	449603	AI655662	Hs.197698	ESTs	5.60
40	414342	AA742181	Hs.75912	KIAA0257 protein	5.59
	422634	NM_016010	Hs.118821	CGI-62 protein	5.56
	435047	AA454985	Hs.54973	cadherin-like protein VR20	5.55
	400268				5.55
	452055	AI377431	Hs.293772	hypothetical protein MGC10858	5.54
45	437073	AI885608	Hs.94122	ESTs	5.54
	434072	H70854	Hs.283059	Homo sapiens PRO1082 mRNA, complete cds	5.53
	418339	AA639902	Hs.104215	ESTs, Moderately similar to SPCN_HUMAN S	5.52
	434551	BE387162	Hs.280858	ESTs, Highly similar to A35661 DNA excis	5.52
	439569	AW602166	Hs.222399	CEGP1 protein	5.51
50	441102	AA973905	Hs.16003	intermediate filament protein syncoilin	5.50
	448310	AI480316		gb:tm26h09.x1 Soares_NFL_T_GBC_S1 Homo s	5.50
	413173	BE076928	Hs.70980	ESTs	5.48
	436246	AW450963	Hs.119991	ESTs	5.48
	449300	AI656959	Hs.222165	ESTs	5.48
55	452823	AB012124	Hs.30696	transcription factor-like 5 (basic helix	5.48
	451403	AA885569	Hs.15727	Homo sapiens cDNA FLJ14511 fis, clone NT	5.46
	417061	AI675944	Hs.188691	Homo sapiens cDNA FLJ12033 fis, clone HE	5.44
	429126	AW172356	Hs.99083	ESTs	5.44
	431316	AA502663	Hs.145037	ESTs	5.44
	439192	AW970536	Hs.105413	ESTs	5.44
60	431938	AA938471	Hs.115242	specific granule protein (28 kDa); cyste	5.44
	451552	AA047233	Hs.33810	ESTs	5.43
	416991	N36389	Hs.295091	KIAA0226 gene product	5.42
	427638	AA406411	Hs.208341	ESTs, Weakly similar to KIAA0989 protein	5.42
	427718	AI798680	Hs.25933	ESTs	5.42
65	438710	AA833907	Hs.178724	ESTs, Weakly similar to ALU1_HUMAN ALU S	5.42
	406076	AL390179	Hs.137011	Homo sapiens mRNA; cDNA DKFZp547P134 (fr	5.40
	431263	AW129203	Hs.13743	ESTs	5.40
	421264	AL039123	Hs.103042	microtubule-associated protein 1B	5.38
	421685	AF189723	Hs.106778	ATPase, Ca++ transporting, type 2C, memb	5.37

	408460	AA054726	Hs.285574	ESTs	5.36
	409091	AW970386	Hs.269423	ESTs	5.36
	421987	AI133161	Hs.286131	CGI-101 protein	5.36
5	428002	AA418703		gb:zv98c03.s1 Soares_NhHMPu_S1 Homo sapi	5.36
	441217	AI922183	Hs.213246	ESTs	5.36
	426006	R49031	Hs.22627	ESTs	5.35
	422806	BE314767	Hs.1581	glutathione S-transferase theta 2	5.34
	432281	AK001239	Hs.274263	hypothetical protein FLJ10377	5.32
10	451982	F13036	Hs.27373	Homo sapiens mRNA; cDNA DKFZp564O1763 (f	5.32
	421129	BE439899	Hs.89271	ESTs	5.31
	444042	NM_004915	Hs.10237	ATP-binding cassette, sub-family G (WHIT	5.31
	410150	AW382942	Hs.6774	ESTs	5.30
	423952	AW877787	Hs.136102	KIAA0853 protein	5.30
15	452822	X85689	Hs.288617	hypothetical protein FLJ22621	5.30
	447752	M73700	Hs.347	lactotransferrin	5.29
	441766	R53790	Hs.23294	hypothetical protein FLJ14393	5.29
	431359	AW993522	Hs.292934	ESTs	5.27
	427212	AW293849	Hs.58279	ESTs, Weakly similar to ALU7_HUMAN ALU S	5.27
20	449916	T60525	Hs.299221	pyruvate dehydrogenase kinase, isoenzyme	5.27
	454014	AW016670	Hs.233275	ESTs	5.27
	419714	AA758751	Hs.98216	ESTs	5.26
	428845	AL157579	Hs.153610	KIAA0751 gene product	5.26
	417333	AL157545	Hs.42179	bromodomain and PHD finger containing, 3	5.24
25	419986	AI345455	Hs.78915	GA-binding protein transcription factor,	5.24
	407182	AA312551	Hs.230157	ESTs	5.22
	420111	AA255652		gb:zs21h11.r1 NCI_CGAP_GCB1 Homo sapiens	5.22
	428058	AI821625	Hs.191602	ESTs	5.22
	459551	AI472808		gb:ij70e07.x1 Soares_NSF_F8_9W_OT_PA_P_S	5.22
30	432524	AI458020	Hs.293287	ESTs	5.22
	436207	AA334774	Hs.12845	hypothetical protein MGC13159	5.22
	410870	U81599	Hs.66731	homeo box B13	5.22
	451418	BE387790	Hs.26369	hypothetical protein FLJ20287	5.22
	409757	NM_001898	Hs.123114	cystatin SN	5.21
35	441124	T97717	Hs.119563	ESTs	5.21
	428593	AW207440	Hs.185973	degenerative spermatocyte (homolog Dros	5.21
	436401	AI087958	Hs.29088	ESTs	5.20
	437113	AA744693		gb:ny26c10.s1 NCI_CGAP_GCB1 Homo sapiens	5.20
40	450947	AI745400	Hs.204662	ESTs	5.20
	453279	AW893940	Hs.59698	ESTs	5.20
	445467	AI239832	Hs.15617	ESTs, Weakly similar to ALU4_HUMAN ALU S	5.19
	448944	AB014605	Hs.22599	atrophin-1 interacting protein 1; activi	5.19
	412198	AA937111	Hs.69165	ESTs	5.18
	422646	H87863	Hs.151380	ESTs, Weakly similar to T16584 hypotheti	5.18
45	438986	AF085888	Hs.269307	ESTs	5.18
	453954	AW118336	Hs.75251	DEAD/H (Asp-Glu-Ala-Asp/His) box binding	5.18
	447541	AK000288	Hs.18800	hypothetical protein FLJ20281	5.18
	434029	AA621763	Hs.170434	Homo sapiens cDNA FLJ14242 fis, clone OV	5.16
	459294	AW977286	Hs.169531	RBP1-like protein	5.16
50	429441	AJ224172	Hs.204096	lipophilin B (uteroglobin family member)	5.16
	424692	AA429834	Hs.151791	KIAA0092 gene product	5.15
	427359	AW020782	Hs.79881	Homo sapiens cDNA: FLJ23006 fis, clone L	5.15
	419872	AI422951	Hs.146162	ESTs	5.15
	429422	AK001494	Hs.202596	Homo sapiens cDNA FLJ10632 fis, clone NT	5.14
55	448902	Z45998	Hs.22543	Homo sapiens mRNA; cDNA DKFZp76111912 (f	5.14
	459055	N23235	Hs.30567	ESTs, Weakly similar to B34087 hypotheti	5.14
	431318	AA502700	Hs.293147	ESTs, Moderately similar to A46010 X-link	5.14
	452953	AI932884	Hs.271741	ESTs, Weakly similar to A46010 X-linked	5.13
	426372	AK000684	Hs.183887	hypothetical protein FLJ22104	5.12
60	434401	AI864131	Hs.71119	Putative prostate cancer tumor suppresso	5.12
	416434	AW163045	Hs.79334	nuclear factor, interleukin 3 regulated	5.11
	410268	AA316181	Hs.61635	six transmembrane epithelial antigen of	5.10
	417517	AF001176	Hs.82238	POP4 (processing of precursor, S. cerev	5.10
	453616	NM_003462	Hs.33846	dynein, axonemal, light intermediate pol	5.10
65	427958	AA418000	Hs.98280	potassium intermediate/small conductance	5.09
	407945	X69208	Hs.606	ATPase, Cu++ transporting, alpha polypep	5.08
	425154	NM_001851	Hs.154850	collagen, type IX, alpha 1	5.08
	412863	AA121673	Hs.59757	zinc finger protein 281	5.06
	420807	AA280627	Hs.57846	ESTs	5.06
	430568	AA769221	Hs.270847	delta-tubulin	5.06

	433687	AA743991		gb.ny57g01.s1 NCI_CGAP_Pr18 Homo sapiens	5.06
	438375	AW015940	Hs.232234	ESTs	5.06
	418092	R45154	Hs.106604	ESTs	5.06
5	418576	AW968159	Hs.289104	Alu-binding protein with zinc finger dom	5.05
	413328	Y15723	Hs.75295	guanylate cyclase 1, soluble, alpha 3	5.04
	414271	AK000275	Hs.75871	protein kinase C binding protein 1	5.04
	432729	AK000292	Hs.278732	hypothetical protein FLJ20285	5.04
	433433	AI692623	Hs.121513	Homo sapiens clone Z'3-1 placenta expres	5.04
	439662	H97552	Hs.269060	ESTs	5.04
10	439743	AL389956	Hs.283858	Homo sapiens mRNA full length insert cDN	5.04
	417511	AL049176	Hs.82223	chordin-like	5.02
	437814	AI088192	Hs.135474	ESTs, Weakly similar to DDX9_HUMAN ATP-D	5.02
	426342	AF093419	Hs.169378	multiple PDZ domain protein	5.02
	429782	NM_005754	Hs.220689	Ras-GTPase-activating protein SH3-domain	5.02
15	429975	AI167145	Hs.165538	ESTs	5.02
	436209	AW850417	Hs.254020	ESTs, Moderately similar to unnamed prot	5.02
	438571	AW020775	Hs.56022	ESTs	5.02
	450223	AA418204	Hs.241493	natural killer-tumor recognition sequenc	5.02
	408267	AW380525	Hs.267705	tubulin-specific chaperone e	5.01
20	417730	Z44761		gb:HSC28F061 normalized infant brain cDN	5.00
	425465	L18964	Hs.1904	protein kinase C, iota	5.00
	430599	NM_004855	Hs.247118	phosphatidylinositol glycan, class B	5.00
	450961	AW978813	Hs.250867	metallothionein 1E (functional)	5.00
	451386	AB029006	Hs.26334	spastic paraplegia 4 (autosomal dominant	5.00
25	420380	AA640891	Hs.102406	ESTs	4.99
	424947	R77952	Hs.239625	ESTs, Weakly similar to alternatively sp	4.99
	442653	BE269247	Hs.170226	gb:601185486F1 NIH_MGC_8 Homo sapiens cD	4.98
	457211	AW972565	Hs.32399	ESTs, Weakly similar to S51797 vasodilat	4.97
	425851	NM_001490	Hs.159642	glucosaminyl (N-acetyl) transferase 1, c	4.97
30	446279	AA490770	Hs.182382	ESTs	4.96
	433377	AI752713	Hs.43845	ESTs	4.96
	450218	R02018	Hs.168640	ankylosis, progressive (mouse) homolog	4.96
	412715	NM_000947	Hs.74519	primase, polypeptide 2A (58kD)	4.94
	448164	R61680	Hs.26904	ESTs, Moderately similar to Z195_HUMAN Z	4.94
35	420121	AW968271	Hs.191534	ESTs, Weakly similar to ALU1_HUMAN ALU S	4.94
	421689	N87820	Hs.106826	KIAA1696 protein	4.93
	445808	AV655234	Hs.298083	ESTs, Moderately similar to PC4259 ferri	4.92
	416533	BE244053	Hs.79362	retinoblastoma-like 2 (p130)	4.92
	418049	AA211467	Hs.190488	Homo sapiens, Similar to nuclear localiz	4.92
40	436039	AW023323	Hs.121070	ESTs	4.92
	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci	4.91
	420324	AF163474	Hs.96744	prostate androgen-regulated transcript 1	4.91
	403047				4.91
	436899	AA764852	Hs.291567	ESTs	4.90
45	431117	AF003522	Hs.250500	delta (Drosophila)-like 1	4.90
	427617	D42063	Hs.179825	RAN binding protein 2	4.88
	428804	AK000713	Hs.193736	hypothetical protein FLJ20706	4.88
	433050	AI093930	Hs.163440	Homo sapiens cDNA: FLJ21000 fis, clone C	4.88
	418575	AA225313	Hs.222886	ESTs, Weakly similar to TRHY_HUMAN TRICH	4.86
50	432615	AA557191	Hs.55028	ESTs, Weakly similar to I54374 gene NF2	4.86
	412652	AI801777	Hs.6774	ESTs	4.86
	432473	AI202703	Hs.152414	ESTs	4.86
	449071	NM_005872	Hs.22960	breast carcinoma amplified sequence 2	4.86
	450654	AJ245587	Hs.25275	Kruppel-type zinc finger protein	4.85
55	418866	T65754	Hs.100489	gb:yc11c07.s1 Stratagene lung (937210) H	4.85
	407596	R86913		gb:yq30f05.r1 Soares fetal liver spleen	4.84
	456516	BE172704	Hs.222746	KIAA1610 protein	4.84
	426501	AW043782	Hs.293616	ESTs	4.84
	448730	AB032983	Hs.21894	KIAA1157 protein	4.84
60	458339	AW976853	Hs.172843	ESTs	4.83
	422083	NM_001141	Hs.111256	arachidonate 15-lipoxygenase, second typ	4.82
	420159	AI572490	Hs.99785	Homo sapiens cDNA: FLJ21245 fis, clone C	4.82
	424103	NM_001918	Hs.139410	dihydrolipoamide branched chain transacy	4.82
	449535	W15267	Hs.23672	low density lipoprotein receptor-related	4.82
65	422048	NM_012445	Hs.288126	spondin 2, extracellular matrix protein	4.82
	416737	AF154335	Hs.79691	LIM domain protein	4.82
	419972	AL041465	Hs.294038	golgin-67	4.81
	420235	AA256758	Hs.31178	ESTs	4.81
	423412	AF109300	Hs.147924	prostate cancer associated protein 5	4.80

	429598	AA811257	Hs.269710	ESTs	4.80
	457114	AI821625	Hs.191602	ESTs	4.80
	421828	AW891965	Hs.289109	histone deacetylase 3	4.79
5	424602	AK002055	Hs.301129	hypothetical protein FLJ11193	4.78
	428364	AA426565	Hs.160541	ESTs, Moderately similar to ALU1_HUMAN A	4.78
	452335	AW188944	Hs.61272	ESTs	4.78
	410765	AI694972	Hs.66180	nucleosome assembly protein 1-like 2	4.77
	421040	AA715026	Hs.135280	ESTs	4.76
10	421518	AI056392	Hs.208819	ESTs	4.76
	452560	BE077084		ESTs	4.76
	409752	AW963990		gb:EST376063 MAGE resequences, MAGH Homo	4.75
	439703	AF086538	Hs.196245	ESTs	4.75
	418836	AI655499	Hs.161712	ESTs	4.74
	450642	R39773	Hs.7130	copine IV	4.74
15	419879	Z17805	Hs.93564	Homer, neuronal immediate early gene, 2	4.74
	411440	AW749402		gb:QV4-BT0383-281299-061-c06 BT0383 Homo	4.74
	450649	NM_001429	Hs.297722	E1A binding protein p300	4.74
	408738	NM_014785	Hs.47313	KIAA0258 gene product	4.73
20	435020	AW505076	Hs.301855	DiGeorge syndrome critical region gene 8	4.72
	411624	BE145964		KIAA0594 protein	4.72
	439360	AA448488	Hs.55346	ribosomal protein L44	4.72
	440491	R35252	Hs.24944	ESTs, Weakly similar to 2109260A B cell	4.72
	442611	BE077155	Hs.177537	hypothetical protein DKFZp761B1514	4.72
25	443555	N71710	Hs.21398	ESTs, Moderately similar to A Chain A, H	4.72
	453800	BE300741	Hs.288416	hypothetical protein FLJ13340	4.72
	457528	AW973791	Hs.292784	ESTs	4.72
	416795	AI497778	Hs.168053	HBV pX associated protein-8	4.71
	407302	R74206	Hs.268755	ESTs, Weakly similar to I78885 serine/th	4.71
	404721				4.70
30	426261	AW242243	Hs.168670	peroxisomal farnesylated protein	4.70
	431924	AK000850	Hs.272203	Homo sapiens cDNA FLJ20843 fis, clone AD	4.70
	435256	AF193766	Hs.13872	cytokine-like protein C17	4.70
	438295	AI394151	Hs.37932	ESTs	4.70
35	442655	AW027457	Hs.30323	ESTs, Weakly similar to B34087 hypotheti	4.70
	415788	AW628686	Hs.78851	KIAA0217 protein	4.69
	442760	BE075297	Hs.10067	ESTs, Weakly similar to A43932 mucin 2 p	4.69
	432432	AA541323	Hs.115831	ESTs	4.68
	454398	AA463437	Hs.11556	Homo sapiens cDNA FLJ12566 fis, clone NT	4.68
40	452741	BE392914	Hs.30503	Homo sapiens cDNA FLJ11344 fis, clone PL	4.67
	424853	BE549737	Hs.132967	Human EST clone 122887 mariner transpos	4.67
	419706	C04649	Hs.77899	tropomyosin 1 (alpha)	4.66
	412088	AI689496	Hs.108932	ESTs	4.65
	416276	U41060	Hs.79136	LIV-1 protein, estrogen regulated	4.64
45	429281	AA830856	Hs.29808	Homo sapiens cDNA: FLJ21122 fis, clone C	4.64
	448207	AI475490	Hs.170577	ESTs	4.64
	408374	AW025430	Hs.155591	forkhead box F1	4.64
	447162	BE328091	Hs.157396	ESTs, Weakly similar to A46010 X-linked	4.64
	451900	AB023199	Hs.27207	KIAA0982 protein	4.63
50	421437	AW821252	Hs.104336	hypothetical protein	4.63
	418624	AI734080	Hs.104211	ESTs	4.63
	426172	AA371307	Hs.125056	ESTs	4.62
	439831	AW136488	Hs.25545	ESTs	4.61
	452994	AW962597	Hs.31305	KIAA1547 protein	4.61
55	457726	AI217477	Hs.194591	ESTs	4.60
	434629	AA789081	Hs.4029	glioma-amplified sequence-41	4.60
	403764				4.58
	410659	AI080175	Hs.68826	ESTs	4.58
	432383	AK000144	Hs.274449	Homo sapiens cDNA FLJ20137 fis, clone CO	4.58
60	451246	AW189232	Hs.39140	cutaneous T-cell lymphoma tumor antigen	4.58
	433234	AB040928	Hs.65366	KIAA1495 protein	4.57
	424983	AI742434	Hs.169911	ESTs	4.56
	437812	AI582291	Hs.16846	ESTs, Weakly similar to O4HUD1 debrisoqu	4.56
	438447	AI082883	Hs.167593	hypothetical protein FLJ13409; KIAA1711	4.55
	434715	BE005346	Hs.116410	ESTs	4.55
65	447673	AI823987	Hs.182285	ESTs	4.54
	408897	N50204	Hs.283709	lipopolysaccharide specific response-7 p	4.54
	436645	AW023424	Hs.156520	ESTs	4.54
	421247	BE391727	Hs.102910	general transcription factor IIH, polype	4.53
	450377	AB033091	Hs.24936	KIAA1265 protein	4.53

	433644	AW342028	Hs.256112	gb:hb75d03.x1 NCI_CGAP_Ut2 Homo sapiens	4.53
	408321	AW405882	Hs.44205	cortistatin	4.53
	439225	AA192669	Hs.45032	ESTs	4.52
	440348	AW015802	Hs.47023	ESTs	4.52
5	446351	AW444551	Hs.258532	x 001 protein	4.52
	451212	AW902672	Hs.287334	ESTs	4.52
	430294	AI538226	Hs.135184	guanine nucleotide binding protein 4	4.52
	435005	U80743	Hs.4316	trinucleotide repeat containing 12	4.52
	448072	AI459306	Hs.24908	ESTs	4.50
10	403721				4.50
	451018	AW965599	Hs.247324	mitochondrial ribosomal protein S14	4.50
	453070	AK001465	Hs.31575	SEC63, endoplasmic reticulum translocon	4.49
	417412	X16896	Hs.82112	interleukin 1 receptor, type I	4.48
	439735	AI635396	Hs.142846	hypothetical protein	4.48
15	435663	AI023707	Hs.134273	ESTs	4.48
	424036	AA770688	Hs.81946	H2A histone family, member L	4.48
	426386	AA748850	Hs.174877	bladder cancer overexpressed protein	4.48
	408622	AA056060	Hs.202577	Homo sapiens cDNA FLJ12166 fis, clone MA	4.47
	444269	AI590346	Hs.146220	ESTs	4.47
20	430187	AI799909	Hs.158989	ESTs	4.46
	427761	AA412205	Hs.140996	ESTs	4.46
	430261	AA305127	Hs.237225	hypothetical protein HT023	4.46
	444169	AV648170	Hs.58756	ESTs	4.44
	430598	AK001764	Hs.247112	hypothetical protein FLJ10902	4.44
25	412903	BE007967	Hs.155795	ESTs	4.44
	417048	AI088775	Hs.55498	geranylgeranyl diphosphate synthase 1	4.44
	442710	AI015631	Hs.23210	ESTs	4.44
	457413	AA743462	Hs.165337	ESTs	4.44
	400303	AA242758	Hs.79136	LIV-1 protein, estrogen regulated	4.42
30	443268	AI800271	Hs.129445	hypothetical protein FLJ12496	4.42
	438209	AL120659	Hs.6111	aryl-hydrocarbon receptor nuclear trans	4.42
	431724	AA514535	Hs.283704	ESTs	4.41
	412280	AW205116	Hs.272814	hypothetical protein DKFZp434E1723	4.40
	440801	AA906366	Hs.190535	ESTs	4.40
35	452959	AI933416	Hs.189674	ESTs	4.40
	453861	AI026838	Hs.30120	ESTs, Weakly similar to NUCL_HUMAN NUCLE	4.40
	417421	AL138201	Hs.82120	nuclear receptor subfamily 4, group A, m	4.40
	447270	AC002551	Hs.331	general transcription factor IIIC, polyp	4.38
	433641	AF080229		gb:Human endogenous retrovirus K clone 1	4.38
40	447078	AW885727	Hs.301570	ESTs	4.38
	424242	AA337476		hypothetical protein MGC13102	4.37
	408170	AW204516	Hs.31835	ESTs	4.36
	448757	AI366784	Hs.48820	TATA box binding protein (TBP)-associate	4.36
	420021	AA252848	Hs.293557	ESTs	4.36
45	449694	AI659790	Hs.253302	ESTs	4.36
	453867	AI929383	Hs.108196	hypothetical protein DKFZp434N185	4.36
	458712	AI347502	Hs.173066	hypothetical protein FLJ20761	4.36
	417251	AW015242	Hs.99488	ESTs, Weakly similar to YK54_YEAST HYPOT	4.35
	434423	NM_006769	Hs.3844	LIM domain only 4	4.35
50	423427	AL137612	Hs.285848	KIAA1454 protein	4.34
	415715	F30364		ESTs	4.33
	404561				4.32
	422969	AA782536	Hs.122647	N-myristoyltransferase 2	4.32
	423685	BE350494	Hs.49753	uveal autoantigen with coiled coil domai	4.32
55	443977	AL120986	Hs.150627	ESTs, Weakly similar to I38022 hypotheti	4.32
	425071	NM_013989	Hs.154424	deiodinase, iodothyronine, type II	4.32
	431583	AL042613	Hs.262476	S-adenosylmethionine decarboxylase 1	4.31
	411379	AI816344	Hs.12554	ESTs, Weakly similar to NPL4_HUMAN NUCLE	4.30
	421476	AW953805	Hs.21887	ESTs	4.30
60	425178	H16097	Hs.161027	ESTs	4.30
	439262	AA832333	Hs.124399	ESTs	4.30
	442818	AK001741	Hs.8739	hypothetical protein FLJ10879	4.30
	421977	W94197	Hs.110165	ribosomal protein L26 homolog	4.29
	437114	AA836641	Hs.163085	ESTs	4.28
65	420195	NA4348	Hs.300794	Homo sapiens cDNA FLJ11177 fis, clone PL	4.28
	418330	BE409405	Hs.94722	ESTs	4.27
	419750	AL079741	Hs.183114	Homo sapiens cDNA FLJ14236 fis, clone NT	4.26
	437065	AL036450	Hs.103238	ESTs	4.26
	455276	BE176479		gb:RC3-HT0585-160300-022-b09 HT0585 Homo	4.24

	416292	AA179233	Hs.42390	nasopharyngeal carcinoma susceptibility	4.24
	423740	Y07701	Hs.132243	aminopeptidase puromycin sensitive	4.24
	442023	AI187878	Hs.144549	ESTs	4.24
5	426764	AA732524	Hs.151464	ESTs, Weakly similar to ALUC_HUMAN !!!!	4.23
	454058	AI273419	Hs.135146	hypothetical protein FLJ13984	4.23
	456511	AA282330	Hs.145668	ESTs	4.22
	448330	AL036449	Hs.207163	ESTs	4.22
	424701	NM_005923	Hs.151988	mitogen-activated protein kinase kinase	4.21
10	432621	AI298501	Hs.12807	ESTs, Weakly similar to T46428 hypotheti	4.20
	445707	AI248720	Hs.114390	ESTs	4.20
	419910	AA662913	Hs.190173	ESTs, Weakly similar to A46010 X-linked	4.20
	424085	NM_002914	Hs.139226	replication factor C (activator 1) 2 (40	4.20
	440749	W22335	Hs.7392	hypothetical protein MGC3199	4.20
	442787	W93048	Hs.227203	hypothetical protein MGC2747	4.20
15	443414	R54594	Hs.25209	ESTs	4.20
	443556	AA256769	Hs.94949	methylmalonyl-CoA epimerase	4.20
	444170	AW613879	Hs.102408	ESTs	4.20
	446751	AA766998	Hs.85874	Human DNA sequence from clone RP11-16L21	4.20
20	421041	N36914	Hs.14691	ESTs, Moderately similar to I38022 hypot	4.19
	447476	BE293466	Hs.20880	ESTs, Weakly similar to I38022 hypotheti	4.19
	448543	AW897741	Hs.21380	Homo sapiens mRNA; cDNA DKFZp586P1124 (f	4.18
	410294	AB014515	Hs.288891	KIAA0615 gene product	4.18
	433607	AA602004	Hs.23260	ESTs	4.18
25	435552	AI686836	Hs.193480	ESTs, Moderately similar to ALU6_HUMAN A	4.18
	447124	AW976438	Hs.17428	RBP1-like protein	4.18
	453308	AW959731	Hs.32538	ESTs	4.17
	439328	W07411	Hs.118212	ESTs, Moderately similar to ALU3_HUMAN A	4.16
	430473	AW130690	Hs.299842	ESTs	4.16
	437257	AI283085	Hs.290931	ESTs, Weakly similar to YFJ7_YEAST HYPOT	4.16
30	438018	AK001160	Hs.5999	hypothetical protein FLJ10298	4.16
	443857	AI089292	Hs.287621	hypothetical protein FLJ14069	4.15
	446711	AF169692	Hs.12450	protocadherin 9	4.15
	419103	Z40229	Hs.96423	hypothetical protein FLJ23033	4.14
	405403				4.14
35	407378	AA299264		ESTs, Moderately similar to I38022 hypot	4.14
	408986	AW298602	Hs.197687	ESTs	4.14
	418727	AA227609	Hs.94834	ESTs	4.14
	434400	AI478211	Hs.186896	Homo sapiens cDNA FLJ11417 fis, clone HE	4.14
	438578	AA811244	Hs.164168	ESTs	4.14
40	450459	AI697193	Hs.299254	Homo sapiens cDNA: FLJ23597 fis, clone L	4.14
	429887	AW366286	Hs.145696	splicing factor (CC1.3)	4.13
	448148	NM_016578	Hs.20509	HBV pX associated protein-8	4.13
	450316	W84446	Hs.17850	hypothetical protein MGC4643	4.12
	417531	NM_003157	Hs.1087	serine/threonine kinase 2	4.12
45	431592	R69016	Hs.293871	hypothetical protein MGC10895s	4.12
	432463	AA548518	Hs.186733	ESTs	4.12
	433613	AA836126	Hs.5669	ESTs	4.12
	434739	AA804487	Hs.144130	ESTs	4.12
	438259	AW205969	Hs.131808	ESTs	4.12
50	425810	AI923627	Hs.31903	ESTs	4.10
	432672	AW973775	Hs.130760	myosin phosphatase, target subunit 2	4.10
	433345	AI681545	Hs.152982	hypothetical protein FLJ13117	4.10
	432712	AB016247	Hs.288031	sterol-C5-desaturase (fungal ERG3, delta	4.09
	453020	AL162039	Hs.31422	Homo sapiens mRNA; cDNA DKFZp434M229 (fr	4.09
55	412045	AA099802	Hs.4299	transmembrane, prostate androgen induced	4.09
	435114	AA775483	Hs.288936	mitochondrial ribosomal protein L9	4.08
	443204	AW205878	Hs.29643	Homo sapiens cDNA FLJ13103 fis, clone NT	4.08
	445459	AI478629	Hs.158465	likely ortholog of mouse putative IKK re	4.08
	438938	H46212	Hs.137221	ESTs	4.07
60	454119	BE549773	Hs.40510	uncoupling protein 4	4.06
	411000	N40449	Hs.201619	ESTs, Weakly similar to S38383 SEB4B pro	4.06
	418926	AA232658	Hs.87070	UDP-glucose:glycoprotein glucosyltransfe	4.06
	424432	AB037821	Hs.146858	protocadherin 10	4.06
	449673	AA002064	Hs.18920	ESTs	4.06
65	429299	AI620463	Hs.99197	hypothetical protein MGC13102	4.06
	422174	AL049325	Hs.112493	Homo sapiens mRNA; cDNA DKFZp564D036 (fr	4.05
	455497	AA112573	Hs.285691	Homo sapiens prostate mRNA, complete cds	4.05
	415138	C18356	Hs.78045	tissue factor pathway inhibitor 2	4.04
	402791				4.04

	426792	AL044854	Hs.172329	KIAA0576 protein	4.04
	438660	U95740	Hs.6349	Homo sapiens, clone IMAGE:3010666, mRNA,	4.04
	442768	AL048534	Hs.48458	ESTs, Weakly similar to ALU8_HUMAN ALU S	4.04
5	447568	AF155655	Hs.18885	CGI-116 protein	4.04
	428342	AI739168	Hs.131798	Homo sapiens cDNA FLJ13458 fis, clone PL	4.04
	453439	AI572438	Hs.32976	guanine nucleotide binding protein 4	4.02
	453857	AL080235	Hs.35861	DKFZP586E1621 protein	4.02
	428249	AA130914	Hs.183291	zinc finger protein 268	4.02
10	432015	AL157504	Hs.159115	Homo sapiens mRNA; cDNA DKFZp586O0724 (f	4.02
	445495	BE622641	Hs.38489	ESTs, Weakly similar to I38022 hypotheti	4.02
	451746	M86178		ESTs	4.02
	452211	AI985513	Hs.233420	ESTs	4.02
	453046	AA284040	Hs.219441	ESTs, Highly similar to CA5B_HUMAN CARBO	4.02
15	456038	AA203285	Hs.294141	ESTs, Weakly similar to alternatively sp	4.02
	452449	AW068658	Hs.20943	ESTs	4.02
	407204	R41933	Hs.140237	ESTs, Weakly similar to ALU1_HUMAN ALU S	4.01
	428046	AW812795	Hs.155381	ESTs, Moderately similar to I38022 hypot	4.01
	438520	AA706319	Hs.98416	ESTs	4.01
	443292	AK000213	Hs.9196	hypothetical protein	4.01
20	432715	AA247152	Hs.200483	ESTs, Weakly similar to KIAA1074 protein	4.00
	403797				4.00
	418347	AA216419	Hs.269295	gb:nc16e03.s1 NCI_CGAP_Pr1 Homo sapiens	4.00
	419459	AW291128	Hs.278422	DKFZP586G1122 protein	4.00
25	420911	U77413	Hs.100293	O-linked N-acetylglucosamine (GlcNAc) tr	4.00
	425176	AW015644	Hs.301430	TEA domain family member 1 (SV40 transcr	4.00
	447505	AL049266	Hs.18724	Homo sapiens mRNA; cDNA DKFZp564F093 (fr	4.00
	453773	AL133761		gb:DKFZp761C1413_r1 761 (synonym: hamy2)	4.00
	434384	AA631910	Hs.162849	ESTs	3.99
30	422471	AA311027	Hs.271894	ESTs, Weakly similar to I38022 hypotheti	3.99
	427386	AW836261	Hs.177486	ESTs	3.98
	433394	AI907753	Hs.93810	cerebral cavernous malformations 1	3.98
	441269	AW015206	Hs.178784	ESTs	3.97
	419629	AB020695	Hs.91662	KIAA0888 protein	3.96
35	435008	AF150262	Hs.162898	ESTs	3.96
	456649	R74441	Hs.117176	poly(A)-binding protein, nuclear 1	3.96
	418723	AA504428	Hs.10487	Homo sapiens, clone IMAGE:3954132, mRNA,	3.96
	428738	NM_000380	Hs.192803	xeroderma pigmentosum, complementation g	3.95
	430456	AA314998	Hs.241503	hypothetical protein	3.95
40	422017	NM_003877	Hs.110776	STAT induced STAT inhibitor-2	3.95
	409960	BE261944	Hs.153028	hexokinase 1	3.95
	455309	AW894017		gb:RC4-NN0027-150400-012-g04 NN0027 Homo	3.95
	450295	AI766732	Hs.201194	ESTs	3.94
	456660	AA909249	Hs.112282	solute carrier family 30 (zinc transport	3.94
45	410908	AA121686	Hs.10592	ESTs	3.94
	447145	AA761073	Hs.192943	TRAF family member-associated NFKB activ	3.94
	449318	AW236021	Hs.108788	Homo sapiens, Similar to RIKEN cDNA 5730	3.94
	449869	W57990	Hs.60059	Homo sapiens cDNA FLJ11478 fis, clone HE	3.94
	411887	AW182924	Hs.128790	ESTs	3.93
50	437531	AI400752	Hs.112259	T cell receptor gamma locus	3.93
	452238	F01811	Hs.187931	ESTs	3.93
	410486	AW235094	Hs.193424	zinc finger protein	3.92
	424882	AI379461	Hs.153636	far upstream element (FUSE) binding prot	3.92
	426269	H15302	Hs.168950	Homo sapiens mRNA; cDNA DKFZp566A1046 (f	3.92
55	427043	AA397679	Hs.298460	ESTs	3.92
	440404	AI015881	Hs.125616	mitochondrial ribosomal protein S5	3.92
	452762	AW501435	Hs.171409	v-akt murine thymoma viral oncogene homo	3.92
	453058	AW612293	Hs.288684	Homo sapiens cDNA FLJ11750 fis, clone HE	3.92
	423583	AL122055	Hs.129836	KIAA1028 protein	3.92
60	408001	AA046458	Hs.95296	ESTs	3.92
	419197	N48921	Hs.27441	KIAA1615 protein	3.91
	428695	AI355647	Hs.189999	purinergic receptor (family A group 5)	3.91
	401747				3.91
	410011	AB020641	Hs.57856	PFTAIR protein kinase 1	3.91
65	432205	AI806583	Hs.125291	ESTs	3.91
	447857	AA081218	Hs.58608	Homo sapiens cDNA FLJ14206 fis, clone NT	3.91
	446494	AA463276	Hs.288906	WW Domain-Containing Gene	3.91
	409928	AL137163	Hs.57549	hypothetical protein dJ473B4	3.90
	411598	BE336654	Hs.70937	H3 histone family, member A	3.90
	424790	AL119344	Hs.13326	ESTs, Weakly similar to 2004399A chromos	3.90

	425707	AF115402	Hs.11713	E74-like factor 5 (ets domain transcript	3.90
	431325	AW026751	Hs.5794	ESTs, Weakly similar to 2109260A B cell	3.89
	451806	NM_003729	Hs.27076	RNA 3'-terminal phosphate cyclase	3.89
5	401045				3.89
	433023	AW864793	Hs.34161	thrombospondin 1	3.89
	452160	BE378541	Hs.279815	cysteine sulfinic acid decarboxylase-rel	3.89
	437372	AA323968	Hs.283631	hypothetical protein DKFZp547G183	3.89
	417067	AJ001417	Hs.81086	solute carrier family 22 (extraneuronal	3.88
	410467	AF102546	Hs.63931	dachshund (Drosophila) homolog	3.88
10	422660	AW297582	Hs.237062	hypothetical protein FLJ22548 similar to	3.88
	431930	AB035301	Hs.272211	cadherin 7, type 2	3.88
	453047	AW023798	Hs.286025	ESTs	3.88
	433891	AA613792		gb:no97h03.s1 NCI_CGAP_Pr2 Homo sapiens	3.88
	401785				3.88
15	431088	AA491824	Hs.196881	ESTs	3.88
	451952	AL120173	Hs.301663	ESTs	3.87
	422089	AA523172	Hs.103135	ESTs, Weakly similar to SFR4_HUMAN SPLIC	3.87
	452277	AL049013	Hs.28783	KIAA1223 protein	3.87
	438279	AA805166	Hs.165165	HIV-1 rev binding protein 2	3.86
20	458229	AI929602	Hs.177	phosphatidylinositol glycan, class H	3.86
	406414				3.86
	417193	AI922189	Hs.288390	hypothetical protein FLJ22795	3.85
	413174	AA723564	Hs.191343	ESTs	3.85
25	433332	AI367347	Hs.127809	Homo sapiens clone TCCCTA00151 mRNA sequ	3.85
	411089	AA456454	Hs.118637	cell division cycle 2-like 1 (PITSLRE pr	3.85
	412494	AL133900	Hs.792	ADP-ribosylation factor domain protein 1	3.84
	413530	AA130158	Hs.19977	ESTs, Moderately similar to ALU8_HUMAN A	3.84
	459592	AL037421	Hs.208746	ESTs, Moderately similar to pot. ORF I [3.84
	418329	AW247430	Hs.84152	cystathionine-beta-synthase	3.83
30	451468	AW503398	Hs.210047	ESTs, Moderately similar to I38022 hypot	3.83
	434804	AA649530		gb:ns44f05.s1 NCI_CGAP_Alv1 Homo sapiens	3.83
	401819				3.82
	424179	F30712		Homo sapiens, clone IMAGE:4285740, mRNA	3.82
35	424850	AA151057	Hs.153498	chromosome 18 open reading frame 1	3.82
	426472	BE246138	Hs.30853	ESTs	3.82
	426625	T78300	Hs.171409	serologically defined colon cancer antig	3.82
	427585	D31152	Hs.179729	collagen, type X, alpha 1 (Schmid metaph	3.82
	427756	AI376540	Hs.15574	ESTs	3.82
	444701	AI916512	Hs.198394	ESTs	3.82
40	423052	M28214	Hs.123072	RAB3B, member RAS oncogene family	3.82
	429259	AA420450	Hs.292911	ESTs, Highly similar to S60712 band-6-pr	3.82
	416111	AA033813	Hs.79018	chromatin assembly factor 1, subunit A (3.82
	433586	T85301		gb:yd78d06.s1 Soares fetal liver spleen	3.81
	438527	AI969251	Hs.143237	RAB7, member RAS oncogene family-like 1	3.81
45	410297	AA148710	Hs.159441	lumican	3.81
	429898	AW117322	Hs.42366	ESTs	3.81
	409079	W87707	Hs.82065	interleukin 6 signal transducer (gp130,	3.80
	419423	D26488	Hs.90315	KIAA0007 protein	3.80
	429643	AA455889	Hs.187548	FYVE-finger-containing Rab5 effector pro	3.80
50	431499	NM_001514	Hs.258561	general transcription factor IIB	3.80
	445060	AA830811	Hs.88808	ESTs	3.80
	449419	R34910	Hs.119172	ESTs	3.80
	450584	AA040403	Hs.60371	ESTs	3.80
	426137	AL040683	Hs.167031	DKFZP566D133 protein	3.79
55	420185	AL044056	Hs.158047	ESTs	3.79
	410076	T05387	Hs.7991	ESTs	3.78
	444078	BE246919	Hs.10290	U5 snRNP-specific 40 kDa protein (hPrp8-	3.78
	417318	AW953937	Hs.12891	ESTs	3.78
	414664	AA587775	Hs.66295	multi-PDZ-domain-containing protein	3.78
60	410275	U85658	Hs.61796	transcription factor AP-2 gamma (activat	3.77
	410503	AW975746	Hs.188662	KIAA1702 protein	3.77
	434170	AA626509	Hs.122329	ESTs	3.77
	421838	AW881089	Hs.108806	Homo sapiens mRNA; cDNA DKFZp566M0947 (f	3.77
	425268	AI807883	Hs.156932	Homo sapiens cDNA FLJ20653 fis, clone KA	3.76
65	431696	AA259068	Hs.267819	protein phosphatase 1, regulatory (inhib	3.76
	411990	AW963624	Hs.31707	ESTs, Weakly similar to YEW4_YEAST HYPOT	3.76
	430291	AV660345	Hs.238126	CGI-49 protein	3.76
	448779	BE042877	Hs.177135	ESTs	3.76
	452682	AA456193	Hs.155606	progesterone membrane binding protein	3.75

5	452598	AI831594	Hs.68647	ESTs, Weakly similar to ALU7_HUMAN ALU S	3.75
	439498	AA908731	Hs.58297	CLLL8 protein	3.75
	440258	AI741633	Hs.125350	ESTs	3.74
	456848	AL121087	Hs.296406	KIAA0685 gene product	3.74
	415082	AA160000	Hs.137396	ESTs, Weakly similar to JC5238 galactosy	3.74
10	420653	AI224532	Hs.88550	ESTs	3.74
	431637	AI879330	Hs.265960	hypothetical protein FLJ10563	3.74
	440411	N30256	Hs.156971	hypothetical protein DKFZp434G1415	3.74
	405917				3.74
	419440	AB020689	Hs.90419	KIAA0882 protein	3.74
15	451230	BE546208	Hs.26090	hypothetical protein FLJ20272	3.73
	429597	NM_003816	Hs.2442	a disintegrin and metalloproteinase doma	3.73
	430144	AI732722	Hs.187694	ERGL protein; ERGIC-53-like protein	3.72
	438394	BE379623	Hs.27693	peptidylprolyl isomerase (cyclophilin)-I	3.72
	440527	AV657117	Hs.184164	ESTs, Moderately similar to S65657 alpha	3.72
20	449433	AI672096	Hs.9012	ESTs, Weakly similar to S26650 DNA-bind	3.72
	456228	BE503227	Hs.134759	ESTs	3.72
	448663	BE614599	Hs.106823	hypothetical protein MGC14797	3.72
	415075	L27479	Hs.77889	Friedreich ataxia region gene X123	3.72
	433544	AI793211	Hs.165372	ESTs, Moderately similar to ALU1_HUMAN A	3.71
25	418293	AI224483	Hs.16063	hypothetical protein FLJ21877	3.71
	449897	AW819642	Hs.24135	transmembrane protein vezatin; hypotheti	3.71
	420297	AI628272	Hs.88323	ESTs, Weakly similar to ALU1_HUMAN ALU S	3.70
	423065	R99158	Hs.194606	Homo sapiens, clone MGC:5406, mRNA, comp	3.70
	429340	N35938	Hs.199429	Homo sapiens mRNA; cDNA DKFZp434M2216 (f	3.70
30	437777	AA768098	Hs.189079	ESTs	3.70
	440351	AF030933	Hs.7179	RAD1 (S. pombe) homolog	3.70
	443603	BE502601	Hs.134289	ESTs, Weakly similar to KIAA1063 protein	3.70
	446965	BE242873	Hs.16677	WD repeat domain 15	3.70
	412350	AI659306	Hs.73826	protein tyrosine phosphatase, non-recept	3.70
35	433852	AI378329	Hs.126629	ESTs	3.70
	433142	AL120697	Hs.110640	ESTs	3.69
	419994	AA282881	Hs.190057	ESTs	3.69
	412628	AI972402	Hs.173902	hypothetical protein MGC2648	3.69
	431416	AA532718	Hs.178604	ESTs	3.69
40	439444	AI277652	Hs.54578	ESTs, Weakly similar to I38022 hypotheti	3.68
	414709	AA704703	Hs.77031	Sp2 transcription factor	3.68
	447397	BE247676	Hs.18442	E-1 enzyme	3.68
	405718				3.68
	425217	AU076696	Hs.155174	CDC5 (cell division cycle 5, S. pombe, h	3.68
45	442242	AV647908	Hs.90424	Homo sapiens cDNA: FLJ23285 fis, clone H	3.68
	424690	BE538356	Hs.151777	eukaryotic translation initiation factor	3.68
	421734	AI318624	Hs.107444	Homo sapiens cDNA FLJ20562 fis, clone KA	3.67
	427221	L15409	Hs.174007	von Hippel-Lindau syndrome	3.67
	439864	AI720078	Hs.291997	ESTs, Weakly similar to A47582 B-cell gr	3.66
50	402408				3.66
	426327	W03242	Hs.44898	Homo sapiens clone TCCCTA00151 mRNA sequ	3.66
	427119	AW880562	Hs.114574	ESTs	3.66
	427356	AW023482	Hs.97849	ESTs	3.66
	452946	X95425	Hs.31092	EphA5	3.66
55	419078	M93119	Hs.89584	insulinoma-associated 1	3.66
	416295	AI064824	Hs.193385	ESTs	3.65
	427144	X95097	Hs.2126	vasoactive intestinal peptide receptor 2	3.65
	447500	AI381900	Hs.159212	ESTs	3.65
	453127	AI696671	Hs.294110	ESTs	3.65
60	423396	AI382555	Hs.127950	bromodomain-containing 1	3.65
	419346	AI830417		polybromo 1	3.64
	441540	C01367	Hs.127128	ESTs	3.64
	446501	AI302616	Hs.150819	ESTs	3.64
	459527	AW977556	Hs.291735	ESTs, Weakly similar to I78885 serine/th	3.63
65	446320	AF126245	Hs.14791	acyl-Coenzyme A dehydrogenase family, me	3.63
	435706	W31254	Hs.7045	GL004 protein	3.63
	400110				3.62
	410313	R10305	Hs.185683	ESTs	3.62
	414713	BE465243	Hs.12664	ESTs	3.62
	436279	AW900372	Hs.180793	ESTs, Weakly similar to S65657 alpha-1C-	3.62
	439818	AL360137	Hs.19934	Homo sapiens mRNA full length insert cDN	3.62
	451797	AW663858	Hs.56120	small inducible cytokine subfamily E, me	3.62
	451294	AI457338	Hs.29894	ESTs	3.62

	434194	AF119847	Hs.283940	Homo sapiens PRO1550 mRNA, partial cds	3.62
	404939				3.62
	408101	AW968504	Hs.123073	CDC2-related protein kinase 7	3.62
5	435846	AA700870	Hs.14304	ESTs	3.61
	432833	N51075	Hs.47191	ESTs	3.61
	427276	AA400269	Hs.49598	ESTs	3.61
	433495	AW373784	Hs.71	alpha-2-glycoprotein 1, zinc	3.60
	403137				3.60
	404165				3.60
10	409571	AA504249	Hs.187585	ESTs	3.60
	410561	BE540255	Hs.6994	Homo sapiens cDNA: FLJ22044 fis, clone H	3.60
	412924	BE018422	Hs.75258	H2A histone family, member Y	3.60
	434228	Z42047	Hs.283978	Homo sapiens PRO2751 mRNA, complete cds	3.60
	436797	AA731491	Hs.178518	hypothetical protein MGC14879	3.60
15	437162	AW005505	Hs.5464	thyroid hormone receptor coactivating pr	3.60
	437444	H46008	Hs.31518	ESTs	3.60
	404210				3.59
	446157	BE270828	Hs.131740	Homo sapiens cDNA: FLJ22562 fis, clone H	3.59
	437587	AI591222	Hs.122421	Human DNA sequence from clone RP1-187J11	3.58
20	423147	AA987927	Hs.131740	Homo sapiens cDNA: FLJ22562 fis, clone H	3.57
	452226	AA024898	Hs.296002	ESTs	3.56
	443775	AF291664	Hs.204732	matrix metalloproteinase 26	3.56
	452501	AB037791	Hs.29716	hypothetical protein FLJ10980	3.56
	428647	AA830050	Hs.124344	ESTs	3.56
25	422443	NM_014707	Hs.116753	histone deacetylase 7B	3.55
	447966	AA340605	Hs.105887	ESTs, Weakly similar to Homolog of rat Z	3.55
	420892	AW975076	Hs.172589	nuclear phosphoprotein similar to S. cer	3.55
	420230	AL034344	Hs.298020	forkhead box C1	3.55
	418428	Y12490	Hs.85092	thyroid hormone receptor interactor 11	3.54
30	428949	AA442153	Hs.104744	hypothetical protein DKFZp434J0617	3.54
	444929	AI685841	Hs.161354	ESTs	3.54
	433339	AF019226	Hs.8036	glioblastoma overexpressed	3.54
	424369	R87622	Hs.26714	KIAA1831 protein	3.54
	433002	AF048730	Hs.279906	cyclin T1	3.53
35	435425	H16263	Hs.31416	ESTs	3.53
	415621	AI648602	Hs.131189	ESTs	3.53
	416974	AF010233	Hs.80667	RALBP1 associated Eps domain containing	3.53
	405793				3.52
	409770	AW499536		gb:U1-HF-BR0p-aji-c-12-0-U1.r1 NIH_MGC_5	3.52
40	425305	AA363025	Hs.155572	Human clone 23801 mRNA sequence	3.52
	428939	AW236550	Hs.131914	ESTs	3.52
	438388	AA806349	Hs.44698	ESTs	3.52
	443703	AV646177	Hs.213021	ESTs	3.52
	457940	AL360159	Hs.30445	Homo sapiens TRlpartite motif protein ps	3.52
45	402444				3.52
	409643	AW450866	Hs.257359	ESTs	3.51
	418250	U29926	Hs.83918	adenosine monophosphate deaminase (isofo	3.51
	432745	AI821926	Hs.269507	gb:nt78f05.x5 NCL_CGAP_Pr3 Homo sapiens	3.51
	414222	AL135173	Hs.878	sorbitol dehydrogenase	3.51
50	430061	AB037817	Hs.230188	KIAA1396 protein	3.51
	421491	H99999	Hs.42736	ESTs	3.50
	422384	AA224077	Hs.42438	Sm protein F	3.50
	434565	T52172		ESTs	3.50
	438379	N23018	Hs.171391	C-terminal binding protein 2	3.50
55	439741	BE379646	Hs.6904	Homo sapiens mRNA full length insert cDN	3.50
	447311	R37010	Hs.33417	Homo sapiens cDNA: FLJ22806 fis, clone K	3.50
	447805	AW627932	Hs.19614	gemin4	3.50
	454265	H03556	Hs.300949	ESTs, Weakly similar to thyroid hormone	3.50
	418838	AW385224	Hs.35198	ectonucleotide pyrophosphatase/phosphodi	3.50
60	448804	AW512213	Hs.42500	ADP-ribosylation factor-like 5	3.50
	409617	BE003760	Hs.55209	Homo sapiens mRNA; cDNA DKFZp434K0514 (f	3.49
	434075	AW003416	Hs.160604	ESTs	3.49
	444190	AI878918	Hs.10526	cysteine and glycine-rich protein 2	3.49
	435017	AA336522	Hs.12854	angiotensin II, type I receptor-associat	3.48
65	423445	NM_014324	Hs.128749	alpha-methylacyl-CoA racemase	3.48
	420271	AI954365	Hs.42892	ESTs	3.48
	443684	AI681307	Hs.166674	ESTs	3.48
	444168	AW379879		gb:RC1-HT0256-081199-011-f01 HT0256 Homo	3.48
	446074	AA079799	Hs.29263	hypothetical protein FLJ11896	3.48

	452582	AL137407	Hs.29911	Homo sapiens mRNA; cDNA DKFZp434M232 (fr	3.48
	431542	H63010	Hs.5740	ESTs	3.48
	432697	AW975050	Hs.293892	ESTs, Weakly similar to ALU4_HUMAN ALU S	3.48
	435572	AW975339	Hs.239828	ESTs, Weakly similar to GAG2_HUMAN RETRO	3.47
5	407192	AA609200		gb:af12e02.s1 Soares_testis_NHT Homo sap	3.47
	413435	X51405	Hs.75360	carboxypeptidase E	3.46
	447210	AF035269	Hs.17752	phosphatidylserine-specific phospholipase	3.46
	447958	AW796524	Hs.68644	Homo sapiens microsomal signal peptidase	3.46
	425312	AA354940	Hs.145958	ESTs	3.46
10	442007	AA301116	Hs.142838	nucleolar phosphoprotein Nopp34	3.46
	417455	AW007066	Hs.18949	ESTs, Weakly similar to CA2B_HUMAN COLLA	3.45
	426931	NM_003416	Hs.2076	zinc finger protein 7 (KOX 4, clone HF.1	3.45
	408739	W01556	Hs.238797	ESTs, Moderately similar to I38022 hypot	3.45
	436024	AI800041	Hs.190555	ESTs	3.45
15	408418	AW963897	Hs.44743	KIAA1435 protein	3.45
	409151	AA306105	Hs.50785	SEC22, vesicle trafficking protein (S. c	3.44
	418626	AW299508	Hs.135230	ESTs	3.44
	420560	AW207748	Hs.59115	ESTs	3.44
	420686	AI950339	Hs.40782	ESTs	3.44
20	428870	AA436831	Hs.36049	ESTs	3.44
	436754	AI061288	Hs.133437	ESTs	3.44
	437960	AI669586	Hs.222194	ESTs	3.44
	452300	AW628045	Hs.28896	Homo sapiens mRNA full length insert cDN	3.44
25	421887	AW161450	Hs.109201	CGI-86 protein	3.44

TABLE 5A shows the accession numbers for those primekeys lacking a unigeneID in Tables 5, 6, and 7. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

10	Pkey: CAT number: Accession:		Unique Eos probeset identifier number Gene cluster number Genbank accession numbers
	Pkey	CAT number	Accession
15	407596	1003489_1	R86913 R86901 H25352 R01370 H43764 AW044451 W21298
	408432	1058667_1	AW195262 R27868 AW811262
	409752	115301_1	AW963990 AA078196 AW749482 AA077468 BE151571 AA376917
	409770	1154048_1	AW499536 AW499553 AW502138 AW499537 AW502136 AW501743
	411440	124577_1	AW749402 AW749403 Z45743 R80376 AA093358
20	411479	1247077_1	AW848047 AW848202 AW848631 AW848142 AW848702 AW848121 AW848632 AW848140 AW848571
			AW848009 AW848067 AW848069 AW848905 AW848214
	411624	1252166_1	BE145964 BE146286 AW854564
	412991	134248_1	AW949013 AA126111
	414269	143133_1	AA298489 AA137165
25	415123	1523390_1	D60925 D60828 D80787
	415715	1548818_1	F30364 F36559 T15435
	416288	1585983_1	H51299 H44619 H46391 R86024 H51892 T72744
	416289	1586037_1	W26333 R05358 H44682
	417730	1695795_1	Z44761 R25801 R11926 R35604
30	418636	177402_1	AW749855 AA225995 AW750208 AW750206
	419346	184129_1	AI830417 AA236612
	419536	185688_1	AA603305 AA244095 AA244183
	420111	190755_1	AA255652 AA280911 AW967920 AA262684
	422219	213547_1	AW978073 AW978072 AA807550 AA306567
35	424179	236389_1	F30712 F35665 AW263888 AI904014 AI904018 AA336927 AA336502
	424242	237181_1	AA337476 AW966227 AA450376 AW960222 AA381051
	428002	285602_1	AA418703 AA418711 BE071915 BE071920 BE071912
	429163	300543_1	AA884766 AW974271 AA592975 AA447312
	432189	342819_1	AA527941 AI810608 AI620190 AA635266
40	432340	345248_1	AA534222 AA632632 T81234
	432363	345469_1	AA534489 AW970240 AW970323
	432966	356839_1	AA650114 AW874148 AA572946
	433586	370470_1	T85301 AW517087 AA601054 BE073959
	433641	37186_1	AF080229 AF080231 AF080230 AF080232 AF080233 AF080234 BE550633 AI636743 AW614951 BE467547
45			AI680833 AI633818 N29986 U87592 U87593 U87590 U87591 S46404 U87587 AA463992 AW206802 AI970376
			AI583718 AI672574 N25695 AW665466 AI818326 AA126128 AI480345 AW013827 AA248638 AI214968
			AA204735 AA207155 AA206262 AA204833 AW003247 AW496808 AI080480 AI631703 AI651023 AI867418
			AW818140 AA502500 AI206199 AI671282 AI352545 BE501030 AI652535 BE465762 AA206331 AW451866
			AA471088 AA206342 AA204834 AA206100 AW021661 AA332922 N66048 AA703396 H92278 AW139734
50			H92683 U87589 U87595 H69001 U87594 BE466420 AI624817 BE466611 AI206344 AA574397 AA348354
			AI493192
	433687	373061_1	AA743991 AA604852 AW272737
	433891	376239_1	AA613792 AW182329 T05304 AW858385
	434415	385931_1	BE177494 AW276909 AA632849
55	434565	38898_1	T52172 AF147324 T52248
	434804	393481_1	AA649530 AA659316 H64973
	437113	433234_1	AA744693 AW750059
	444168	593829_1	AW379879 AI126285 H12014
	448212	755099_1	AI475858 AW969013
	448310	757918_1	AI480316 AW847535
	451746	883303_1	M86178 AI813822 D56993

5	<p>452560 922216_1</p> <p>452712 928309_1</p> <p>453773 980699_1</p> <p>455276 1272541_1</p> <p>455309 1278153_1</p>	<p>BE077084 AW139963 AW863127 AW806209 AW806204 AW806205 AW806206 AW806211 AW806212</p> <p>AW806207 AW806208 AW806210 AI907497</p> <p>AW838616 AW838660 BE144343 AI914520 AW888910 BE184854 BE184784</p> <p>AL133761 AL133767</p> <p>BE176479 BE176678 BE176357 BE176550 AW886079 BE176676 BE176615 BE176555 BE176489 BE176610</p> <p>BE176362</p> <p>AW894017 AW893956 AW894032</p>
---	---	--

TABLE 5B shows the genomic positioning for those primekeys lacking unigene ID's and accession numbers in Tables 5, 6, and 7. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

10	Pkey:	Unique number corresponding to an Eos probeset		
	Ref:	Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.		
	Strand:	Indicates DNA strand from which exons were predicted.		
	Nt_position:	Indicates nucleotide positions of predicted exons.		
15				
	Pkey	Ref	Strand	Nt_position
20	401045	8117619	Plus	90044-90184,91111-91345
	401424	8176894	Plus	24223-24428
	401451	6634068	Minus	119926-121272
	401714	6715702	Plus	96484-96681
	401747	9789672	Minus	118596-118816,119119-119244,119609-119761,120422-120990,130161-130381,130468-130593,131097-131258,131866-131932,132451-132575,133580-134011
25	401785	7249190	Minus	165776-165996,166189-166314,166408-166569,167112-167268,167387-167469,168634-168942
	401819	7467933	Minus	28217-28486
	402408	9796239	Minus	110326-110491
	402444	9796614	Plus	28391-28517
	402791	6137008	Minus	51036-51207
30	403047	3540153	Minus	59793-59968
	403137	9211494	Minus	92349-92572,92958-93084,93579-93712,93949-94072,94591-94748,95214-95337
	403721	7528046	Minus	156647-157366
	403764	7717105	Minus	118692-118853
	403797	8099896	Minus	123065-125008
35	404165	9926489	Minus	69025-69128
	404210	5006246	Plus	169926-170121
	404253	9367202	Minus	55675-56055
	404561	9795980	Minus	69039-70100
	404571	7249169	Minus	112450-112648
40	404721	9856648	Minus	173763-174294
	404915	7341766	Minus	100915-101087
	404939	6862697	Plus	175318-175476
	405403	6850244	Minus	37491-37670,40951-41031
	405685	4508129	Minus	37956-38097
45	405718	9795467	Plus	113080-113266
	405793	1405887	Minus	89197-89453
	405876	6758747	Plus	39694-40031
	405917	7712162	Minus	106829-107213
	406414	9256407	Plus	49593-49850
50	406554	7711566	Plus	106956-107121

TABLE 6:286 GENES ENCODING EXTRACELLULAR OR CELL SURFACE PROTEINS UP-REGULATED IN PROSTATE CANCER COMPARED TO NORMAL ADULT TISSUES

Table 6 shows 286 genes up-regulated in prostate cancer compared to normal adult tissues that are likely to be extracellular or cell-surface proteins. These were selected as for Table 5 and the predicted protein contained a structural domain that is indicative of extracellular localization (e.g. egf, 7tm domains).

	Pkey: ExAccn: UnigeneID: Unigene Title: R1:	Unique Eos probeset identifier number Exemplar Accession number, Genbank accession number Unigene number Unigene gene title Ratio of tumor to normal tissue			
10	Pkey	ExAccn	UnigeneID	Unigene Title	R1
	409361	NM_005982	Hs.54416	sine oculis homeobox (Drosophila) homolo	48.28
	409731	AA125985	Hs.56145	thymosin, beta, identified in neuroblast	45.24
	400298	AA032279	Hs.61635	six transmembrane epithelial antigen of	43.48
15	420154	AI093155	Hs.95420	JM27 protein	41.12
	426747	AA535210	Hs.171995	kallikrein 3, (prostate specific antigen	31.80
	400299	X07730	Hs.171995	kallikrein 3, (prostate specific antigen	24.91
	425075	AA506324	Hs.1852	acid phosphatase, prostate	24.23
	424846	AU077324	Hs.1832	neuropeptide Y	23.57
20	405685				20.90
	420757	X78592	Hs.99915	androgen receptor (dihydrotestosterone r	19.72
	418994	AA296520	Hs.89546	selectin E (endothelial adhesion molecul	19.56
	452792	AB037765	Hs.30652	KIAA1344 protein	17.39
	445472	AB006631	Hs.12784	Homo sapiens mRNA for KIAA0293 gene, par	17.00
25	414565	AA502972	Hs.183390	hypothetical protein FLJ13590	16.82
	431716	D89053	Hs.268012	fatty-acid-Coenzyme A ligase, long-chain	16.60
	408430	S79876	Hs.44926	dipeptidylpeptidase IV (CD26, adenosine	16.28
	408000	L11690	Hs.620	bullous pemphigoid antigen 1 (230/240kD)	15.54
	430226	BE245562	Hs.2551	adrenergic, beta-2-, receptor, surface	15.40
30	444484	AK002126	Hs.11260	hypothetical protein FLJ11264	14.76
	418601	AA279490	Hs.86368	calmegin	14.56
	448999	AF179274	Hs.22791	transmembrane protein with EGF-like and	14.55
	416182	NM_004354	Hs.79069	cyclin G2	12.94
	420544	AA677577	Hs.98732	Homo sapiens Chromosome 16 BAC clone CIT	12.79
35	445413	AA151342	Hs.12677	CGI-147 protein	12.64
	453930	AA419466	Hs.36727	hypothetical protein FLJ10903	12.22
	440286	U29589	Hs.7138	cholinergic receptor, muscarinic 3	12.04
	452784	BE463857	Hs.151258	hypothetical protein FLJ21062	11.86
	450203	AF097994	Hs.301528	L-kynurenine/alpha-aminoadipate aminotra	11.68
40	448045	AJ297436	Hs.20166	prostate stem cell antigen	11.51
	449650	AF055575	Hs.23838	calcium channel, voltage-dependent, L ty	11.18
	420381	D50640	Hs.337616	phosphodiesterase 3B, cGMP-inhibited	11.10
	425665	AK001050	Hs.159066	hypothetical protein FLJ10188	11.08
	425710	AF030880	Hs.159275	solute carrier family, member 4	11.08
45	428728	NM_016625	Hs.191381	hypothetical protein	11.04
	407021	U52077		gb:Human mariner1 transposase gene, comp	11.02
	410733	D84284	Hs.66052	CD38 antigen (p45)	11.02
	452340	NM_002202	Hs.505	ISL1 transcription factor, LIM/homeodoma	10.85
	428819	AL135623	Hs.193914	KIAA0575 gene product	10.48
50	421991	NM_014918	Hs.110488	KIAA0990 protein	10.04
	431217	NM_013427	Hs.250830	Rho GTPase activating protein 6	9.75
	421470	R27496	Hs.1378	annexin A3	9.64
	409262	AK000631	Hs.52256	hypothetical protein FLJ20624	9.45
	435980	AF274571	Hs.129142	deoxyribonuclease II beta	9.24
55	421246	AW582962	Hs.102897	CGI-47 protein	9.20
	410001	AB041036	Hs.57771	kallikrein 11	9.03
	441791	AW372449	Hs.175982	hypothetical protein FLJ21159	9.02

	404571				8.66
	456497	AW967956	Hs.123648	ESTs, Weakly similar to AF108460 1 ubinu	8.56
	419968	X04430	Hs.93913	interleukin 6 (interferon, beta 2)	8.36
5	433172	AB037841	Hs.102652	hypothetical protein ASH1	8.30
	422631	BE218919	Hs.118793	hypothetical protein FLJ10688	8.27
	427674	NM_003528	Hs.2178	H2B histone family, member Q	8.20
	404915				8.08
	452259	AA317439	Hs.28707	signal sequence receptor, gamma (translo	8.06
10	452891	N75582	Hs.212875	ESTs, Weakly similar to DYH9_HUMAN CILIA	8.02
	439731	AI953135	Hs.45140	hypothetical protein FLJ14084	7.98
	419839	U24577	Hs.93304	phospholipase A2, group VII (platelet-ac	7.68
	420120	AL049610	Hs.95243	transcription elongation factor A (SII)-	7.64
	424099	AF071202	Hs.139336	ATP-binding cassette, sub-family C (CFTR	7.64
15	448706	AW291095	Hs.21814	interleukin 20 receptor, alpha	7.52
	410227	AB009284	Hs.61152	exostoses (multiple)-like 2	7.49
	425211	M18667	Hs.1867	progastricsin (pepsinogen C)	7.35
	441736	AW292779	Hs.169799	ESTs	7.28
	419991	AJ000098	Hs.94210	eyes absent (Drosophila) homolog 1	7.20
20	425018	BE245277	Hs.154196	E4F transcription factor 1	7.20
	424560	AA158727	Hs.150555	protein predicted by clone 23733	7.18
	409110	AA191493	Hs.48778	niban protein	7.10
	421566	NM_000399	Hs.1395	early growth response 2 (Krox-20 (Drosop	7.04
	431725	X65724	Hs.2839	Norrie disease (pseudoglioma)	6.98
25	425782	U66468	Hs.159525	cell growth regulatory with EF-hand doma	6.85
	427408	AA583206	Hs.2156	RAR-related orphan receptor A	6.79
	435604	AA625279	Hs.26892	uncharacterized bone marrow protein BM04	6.73
	415874	AF091622	Hs.78893	KIAA0244 protein	6.54
	401451				6.52
30	431778	AL080276	Hs.268562	regulator of G-protein signalling 17	6.51
	409089	NM_014781	Hs.50421	KIAA0203 gene product	6.50
	431992	NM_002742	Hs.2891	protein kinase C, mu	6.49
	404253				6.42
	421552	AF026692	Hs.105700	secreted frizzled-related protein 4	6.41
35	416806	NM_000288	Hs.79993	peroxisomal biogenesis factor 7	6.38
	431958	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta	6.30
	439366	AF100143	Hs.6540	fibroblast growth factor 13	6.30
	416836	D54745	Hs.80247	cholecystokinin	6.30
	433383	AF034837	Hs.192731	double-stranded RNA specific adenosine d	6.29
	450728	AW162923	Hs.25363	presenilin 2 (Alzheimer disease 4)	6.25
40	413384	NM_000401	Hs.75334	exostoses (multiple) 2	6.22
	423349	AF010258	Hs.127428	homeo box A9	6.20
	424800	AL035588	Hs.153203	MyoD family inhibitor	6.18
	425451	AF242769	Hs.157461	mesenchymal stem cell protein DSC54	6.14
45	447359	NM_012093	Hs.18268	adenylate kinase 5	6.00
	410889	X91662	Hs.66744	twist (Drosophila) homolog (acrocephalos	5.97
	408829	NM_006042	Hs.48384	heparan sulfate (glucosamine) 3-O-sulfot	5.94
	453911	AW503857	Hs.4007	Sarcolemmal-associated protein	5.94
	408875	NM_015434	Hs.48604	DKFZP434B168 protein	5.92
50	450480	X82125	Hs.25040	zinc finger protein 239	5.90
	451684	AF216751	Hs.26813	CDA14	5.88
	400301	X03635	Hs.1657	estrogen receptor 1	5.78
	415077	L41607	Hs.934	glucosaminyl (N-acetyl) transferase 2, I	5.74
	418852	BE537037	Hs.273294	hypothetical protein FLJ20069	5.72
55	446867	AB007891	Hs.16349	KIAA0431 protein	5.72
	410232	AW372451	Hs.61184	CGI-79 protein	5.70
	422762	AL031320	Hs.119976	Human DNA sequence from clone RP1-20N2 o	5.70
	450616	AL133067	Hs.302689	hypothetical protein	5.70
	408621	AI970672	Hs.46638	chromosome 11 open reading frame 8	5.65
60	439671	AW162840	Hs.6641	kinesin family member 5C	5.64
	410196	AI936442	Hs.59838	hypothetical protein FLJ10808	5.60
	429170	NM_001394	Hs.2359	dual specificity phosphatase 4	5.60
	440738	AI004650	Hs.225674	WD repeat domain 9	5.60
	414342	AA742181	Hs.75912	KIAA0257 protein	5.59
65	422634	NM_016010	Hs.118821	CGI-62 protein	5.56
	400268				5.55
	439569	AW602166	Hs.222399	CEGP1 protein	5.51
	452823	AB012124	Hs.30696	transcription factor-like 5 (basic helix	5.48
	431938	AA938471	Hs.54431	specific granule protein (28 kDa); cyste	5.44
	427638	AA406411	Hs.208341	ESTs, Weakly similar to KIAA0989 protein	5.42

5	421264	AL039123	Hs.103042	microtubule-associated protein 1B	5.38
	421685	AF189723	Hs.106778	ATPase, Ca++ transporting, type 2C, memb	5.37
	421987	AI133161	Hs.286131	CGI-101 protein	5.36
	422806	BE314767	Hs.1581	glutathione S-transferase theta 2	5.34
	432281	AK001239	Hs.274263	hypothetical protein FLJ10377	5.32
10	451982	F13036	Hs.27373	Homo sapiens mRNA; cDNA DKFZp564O1763 (f	5.32
	444042	NM_004915	Hs.10237	ATP-binding cassette, sub-family G (WHIT	5.31
	447752	M73700	Hs.105938	lactotransferrin	5.29
	451418	BE387790	Hs.26369	hypothetical protein FLJ20287	5.22
	428593	AW207440	Hs.185973	degenerative spermatocyte (homolog Droso	5.21
15	447541	AK000288	Hs.18800	hypothetical protein FLJ20281	5.18
	459294	AW977286	Hs.17428	RBP1-like protein	5.16
	424692	AA429834	Hs.151791	KIAA0092 gene product	5.15
	416434	AW163045	Hs.79334	nuclear factor, interleukin 3 regulated	5.11
	410268	AA316181	Hs.61635	six transmembrane epithelial antigen of	5.10
20	417517	AF001176	Hs.82238	POP4 (processing of precursor, S. cerev	5.10
	453616	NM_003462	Hs.33846	dynein, axonemal, light intermediate pol	5.10
	427958	AA418000	Hs.98280	potassium intermediate/small conductance	5.09
	407945	X69208	Hs.606	ATPase, Cu++ transporting, alpha polypep	5.08
	418576	AW968159	Hs.289104	Alu-binding protein with zinc finger dom	5.05
25	413328	Y15723	Hs.75295	guanylate cyclase 1, soluble, alpha 3	5.04
	432729	AK000292	Hs.278732	hypothetical protein FLJ20285	5.04
	426342	AF093419	Hs.169378	multiple PDZ domain protein	5.02
	429782	NM_005754	Hs.220689	Ras-GTPase-activating protein SH3-domain	5.02
	436209	AW850417	Hs.254020	ESTs, Moderately similar to unnamed prot	5.02
30	430599	NM_004855	Hs.247118	phosphatidylinositol glycan, class B	5.00
	451386	AB029006	Hs.26334	spastic paraplegia 4 (autosomal dominant	5.00
	457211	AW972565	Hs.32399	ESTs, Weakly similar to S51797 vasodilat	4.97
	425851	NM_001490	Hs.159642	glucosaminyl (N-acetyl) transferase 1, c	4.97
	421689	N87820	Hs.106826	KIAA1696 protein	4.93
35	416533	BE244053	Hs.79362	retinoblastoma-like 2 (p130)	4.92
	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci	4.91
	403047				4.91
	431117	AF003522	Hs.250500	delta (Drosophila)-like 1	4.90
	427617	D42063	Hs.199179	RAN binding protein 2	4.88
40	428804	AK000713	Hs.193736	hypothetical protein FLJ20706	4.88
	449071	NM_005872	Hs.22960	breast carcinoma amplified sequence 2	4.86
	407596	R86913		gb:yq30f05.r1 Soares fetal liver spleen	4.84
	456516	BE172704	Hs.222746	KIAA1610 protein	4.84
	458339	AW976853	Hs.172843	ESTs	4.83
45	422083	NM_001141	Hs.111256	arachidonate 15-lipoxygenase, second typ	4.82
	449535	W15267	Hs.23672	low density lipoprotein receptor-related	4.82
	422048	NM_012445	Hs.288126	spondin 2, extracellular matrix protein	4.82
	424602	AK002055	Hs.151046	hypothetical protein FLJ11193	4.78
	410765	AI694972	Hs.66180	nucleosome assembly protein 1-like 2	4.77
50	419879	Z17805	Hs.93564	Homer, neuronal immediate early gene, 2	4.74
	450649	NM_001429	Hs.25272	E1A binding protein p300	4.74
	411624	BE145964	Hs.103283	KIAA0594 protein	4.72
	404721				4.70
	426261	AW242243	Hs.168670	peroxisomal farnesylated protein	4.70
55	416276	U41060	Hs.79136	LIV-1 protein, estrogen regulated	4.64
	408374	AW025430	Hs.155591	forkhead box F1	4.64
	451900	AB023199	Hs.27207	KIAA0982 protein	4.63
	421437	AW821252	Hs.104336	hypothetical protein	4.63
	434629	AA789081	Hs.4029	glioma-amplified sequence-41	4.60
60	403764				4.58
	421247	BE391727	Hs.102910	general transcription factor IIH, polype	4.53
	403721				4.50
	453070	AK001465	Hs.31575	SEC63, endoplasmic reticulum translocon	4.49
	417412	X16896	Hs.82112	interleukin 1 receptor, type I	4.48
65	439735	AI635386	Hs.142846	hypothetical protein	4.48
	430261	AA305127	Hs.237225	hypothetical protein HT023	4.46
	430598	AK001764	Hs.247112	hypothetical protein FLJ10902	4.44
	400303	AA242758	Hs.79136	LIV-1 protein, estrogen regulated	4.42
	438209	AL120659	Hs.6111	aryl-hydrocarbon receptor nuclear trans	4.42
	417421	AL138201	Hs.82120	nuclear receptor subfamily 4, group A, m	4.40
	447270	AC002551	Hs.331	general transcription factor IIIC, polyp	4.38
	434423	NM_006769	Hs.3844	LIM domain only 4	4.35
	404561				4.32

5	422969	AA782536	Hs.122647	N-myristoyltransferase 2	4.32
	423685	BE350494	Hs.49753	uvcal autoantigen with coiled coil domai	4.32
	425071	NM_013989	Hs.154424	deiodinase, iodothyronine, type II	4.32
	431583	AL042613	Hs.262476	S-adenosylmethionine decarboxylase 1	4.31
	442818	AK001741	Hs.8739	hypothetical protein FLJ10879	4.30
10	423740	Y07701	Hs.293007	aminopeptidase puromycin sensitive	4.24
	424701	NM_005923	Hs.151988	mitogen-activated protein kinase kinase	4.21
	424085	NM_002914	Hs.139226	replication factor C (activator 1) 2 (40	4.20
	410294	AB014515	Hs.323712	KIAA0615 gene product	4.18
	447124	AW976438	Hs.17428	RBP1-like protein	4.18
15	438018	AK001160	Hs.5999	hypothetical protein FLJ10298	4.16
	443857	AI089292	Hs.287621	hypothetical protein FLJ14069	4.15
	446711	AF169692	Hs.12450	protocadherin 9	4.15
	405403				4.14
	448148	NM_016578	Hs.20509	HBV pX associated protein-8	4.13
20	417531	NM_003157	Hs.1087	serine/threonine kinase 2	4.12
	433345	AI681545	Hs.152982	hypothetical protein FLJ13117	4.10
	432712	AB016247	Hs.288031	sterol-C5-desaturase (fungal ERG3, delta	4.09
	435114	AA775483	Hs.288936	mitochondrial ribosomal protein L9	4.08
	445459	AI478629	Hs.158465	likely ortholog of mouse putative IKK re	4.08
25	402791				4.04
	438660	U95740	Hs.6349	Homo sapiens, clone IMAGE:3010666, mRNA,	4.04
	447568	AF155655	Hs.18885	CGI-116 protein	4.04
	452211	AI985513	Hs.233420	ESTs	4.02
	443292	AK000213	Hs.9196	hypothetical protein	4.01
30	420911	U77413	Hs.100293	O-linked N-acetylglucosamine (GlcNAc) tr	4.00
	428738	NM_000380	Hs.192803	xeroderma pigmentosum, complementation g	3.95
	430456	AA314998	Hs.241503	hypothetical protein	3.95
	437531	AI400752	Hs.112259	T cell receptor gamma locus	3.93
	428695	AI355647	Hs.189999	purinergic receptor (family A group 5)	3.91
35	410011	AB020641	Hs.57856	PFTAIRE protein kinase 1	3.91
	446494	AA463276	Hs.288906	WW Domain-Containing Gene	3.91
	409928	AL137163	Hs.57549	hypothetical protein dJ473B4	3.90
	411598	BE336654	Hs.70937	H3 histone family, member A	3.90
	425707	AF115402	Hs.11713	E74-like factor 5 (ets domain transcript	3.90
40	451806	NM_003729	Hs.27076	RNA 3'-terminal phosphate cyclase	3.89
	401045				3.89
	437372	AA323968	Hs.283631	hypothetical protein DKFZp547G183	3.89
	417067	AJ001417	Hs.81086	solute carrier family 22 (extraneuronal	3.88
	410467	AF102546	Hs.63931	dachshund (Drosophila) homolog	3.88
45	431930	AB035301	Hs.272211	cadherin 7, type 2	3.88
	453047	AW023798	Hs.286025	ESTs	3.88
	401785				3.88
	458229	AI929602	Hs.177	phosphatidylinositol glycan, class H	3.86
	406414				3.86
50	412494	AL133900	Hs.792	ADP-ribosylation factor domain protein 1	3.84
	418329	AW247430	Hs.84152	cystathionine-beta-synthase	3.83
	424850	AA151057	Hs.153498	chromosome 18 open reading frame 1	3.82
	427585	D31152	Hs.179729	collagen, type X, alpha 1 (Schmid metaph	3.82
	423052	M28214	Hs.123072	RAB3B, member RAS oncogene family	3.82
55	416111	AA033813	Hs.79018	chromatin assembly factor 1, subunit A (3.82
	419423	D26488	Hs.90315	KIAA0007 protein	3.80
	429643	AA455889	Hs.167279	FYVE-finger-containing Rab5 effector pro	3.80
	431499	NM_001514	Hs.258561	general transcription factor IIB	3.80
	444078	BE246919	Hs.10290	U5 snRNP-specific 40 kDa protein (hPrp8-	3.78
60	430291	AV660345	Hs.238126	CGI-49 protein	3.76
	431637	AI879330	Hs.265960	hypothetical protein FLJ10563	3.74
	440411	N30256	Hs.151093	hypothetical protein DKFZp434G1415	3.74
	405917				3.74
	451230	BE546208	Hs.26090	hypothetical protein FLJ20272	3.73
65	429597	NM_003816	Hs.2442	a disintegrin and metalloproteinase doma	3.73
	415075	L27479	Hs.77889	Friedreich ataxia region gene X123	3.72
	440351	AF030933	Hs.7179	RAD1 (S. pombe) homolog	3.70
	443603	BE502601	Hs.134289	ESTs, Weakly similar to KIAA1063 protein	3.70
	446965	BE242873	Hs.16677	WD repeat domain 15	3.70
	412350	AI659306	Hs.73826	protein tyrosine phosphatase, non-recept	3.70
	433852	AI378329	Hs.126629	ESTs	3.70
	447397	BE247676	Hs.18442	E-1 enzyme	3.68
	405718				3.68

	425217	AU076696	Hs.155174	ODC5 (cell division cycle 5, S. pombe, h	3.68
	421734	AI318624	Hs.107444	Homo sapiens cDNA FLJ20562 fis, clone KA	3.67
	427221	L15409	Hs.174007	von Hippel-Lindau syndrome	3.67
	402408				3.66
5	452946	X95425	Hs.31092	EphA5	3.66
	419078	M93119	Hs.89584	insulinoma-associated 1	3.66
	427144	X95097	Hs.2126	vasoactive intestinal peptide receptor 2	3.65
	423396	AI382555	Hs.127950	bromodomain-containing 1	3.65
10	446320	AF126245	Hs.14791	acyl-Coenzyme A dehydrogenase family, me	3.63
	404939				3.62
	403137				3.60
	437162	AW005505	Hs.5464	thyroid hormone receptor coactivating pr	3.60
	404210				3.59
	443775	AF291664	Hs.204732	matrix metalloproteinase 26	3.56
15	452501	AB037791	Hs.29716	hypothetical protein FLJ10980	3.56
	422443	NM_014707	Hs.116753	histone deacetylase 7B	3.55
	420230	AL034344	Hs.284186	forkhead box C1	3.55
	418428	Y12490	Hs.85092	thyroid hormone receptor interactor 11	3.54
	433002	AF048730	Hs.279906	cyclin T1	3.53
20	405793				3.52
	457940	AL360159	Hs.306517	Homo sapiens TRlpartite motif protein ps	3.52
	402444				3.52
	418250	U29926	Hs.83918	adenosine monophosphate deaminase (isofo	3.51
	414222	AL135173	Hs.878	sorbitol dehydrogenase	3.51
25	422384	AA224077	Hs.42438	Sm protein F	3.50
	447805	AW627932	Hs.19614	gemin4	3.50
	454265	H03556	Hs.300949	ESTs, Weakly similar to thyroid hormone	3.50
	423445	NM_014324	Hs.128749	alpha-methylacyl-CoA racemase	3.48
	413435	X51405	Hs.75360	carboxypeptidase E	3.46
30	447210	AF035269	Hs.17752	phosphatidylserine-specific phospholipas	3.46
	426931	NM_003416	Hs.2076	zinc finger protein 7 (KOX 4, clone HF.1	3.45
	408418	AW963897	Hs.44743	KIAA1435 protein	3.45
	421887	AW161450	Hs.109201	CGI-86 protein	3.44

Table 7: 42 GENES ENCODING SMALL MOLECULE TARGETS UP-REGULATED IN PROSTATE CANCER COMPARED TO NORMAL ADULT TISSUES

Table 7 shows 42 genes up-regulated in prostate cancer compared to normal adult tissues that are likely to be small molecule targets. These were selected as for Table 5 and the predicted protein contained a structural domain that is indicative of a drugable structure (e.g. protease, kinase, phosphatase, receptor). The functional domain is indicated for each gene.

10 Pkey: Unique Eos probeset identifier number
 ExAccn: Exemplar Accession number, Genbank accession number
 UnigenelD: Unigene number
 Unigene Title: Unigene gene title
 PSDomain: Protein Structural Domain
 15 R1: Ratio of tumor vs. normal tissue

	Pkey	ExAccn	UnigenelD	Unigene Title	PSDomain	R1
20	426747	AA535210	Hs.171995	kallikrein 3, (prostate specific antigen	trypsin	31.80
	400299	X07730	Hs.171995	kallikrein 3, (prostate specific antigen	trypsin	24.91
	420757	X78592	Hs.99915	androgen receptor (dihydrotestosterone r	Androgen_recep,hormone_rec,zf-C4	19.72
	408430	S79876	Hs.44926	dipeptidylpeptidase IV (CD26, adenosine	DPPIV_N_term,Peptidase_S9	16.28
	430226	BE245562	Hs.2551	adrenergic, beta-2-, receptor, surface	7tm_1	15.40
25	411096	U80034	Hs.68583	mitochondrial intermediate peptidase	Peptidase_M3	14.81
	440286	U29589	Hs.7138	cholinergic receptor, muscarinic 3	7tm_1	12.04
	420381	D50640	Hs.337616	phosphodiesterase 3B, cGMP-inhibited	PDEase	11.10
	407021	U52077		gb:Human mariner1 transposase gene, comp	SET,Transposase_1	11.02
	401424				arginase	9.58
30	410001	AB041036	Hs.57771	kallikrein 11	trypsin	9.03
	428330	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin,	Peptidase_M10	8.76
	424099	AF071202	Hs.139336	ATP-binding cassette, sub-family C (CFTR	ABC_tran,ABC_membrane	7.64
	419991	AJ000098	Hs.94210	eyes absent (Drosophila) homolog 1	Hydrolase	7.20
	431992	NM_002742	Hs.2891	protein kinase C, mu	pkinase,DAG_PE-bind,PH	6.49
35	447359	NM_012093	Hs.18268	adenylate kinase 5	adenylatekinase	6.00
	400301	X03635	Hs.1657	estrogen receptor 1	Oest_recep,zf-C4,hormone_rec	5.78
	421685	AF189723	Hs.106778	ATPase, Ca++ transporting, type 2C, memb	E1-E2_ATPase,Hydrolase	5.37
	444042	NM_004915	Hs.10237	ATP-binding cassette, sub-family G (WHIT	ABC_tran	5.31
	447752	M73700	Hs.105938	lactotransferrin	transferrin,7tm_1	5.29
40	407945	X69208	Hs.606	ATPase, Cu++ transporting, alpha polypep	E1-E2_ATPase,Hydrolase,HMA	5.08
	403047				trypsin	4.91
	427617	D42063	Hs.199179	RAN binding protein 2	Ran_BP1,zf-RanBP,TPR,pro_isomerase	4.88
	422083	NM_001141	Hs.111256	arachidonate 15-lipoxygenase, second typ	lipoxygenase,PLAT	4.82
	449535	W15267	Hs.23672	low density lipoprotein receptor-related	ldl_recept_b,ldl_recept_a,EGF	4.82
45	425071	NM_013989	Hs.154424	deiodinase, iodothyronine, type II	T4_deiodinase	4.32
	423740	Y07701	Hs.293007	aminopeptidase puromycin sensitive	Peptidase_M1	4.24
	424701	NM_005923	Hs.151988	mitogen-activated protein kinase kinase	pkinase	4.21
	424085	NM_002914	Hs.139226	replication factor C (activator 1) 2 (40	AAA,Viral_helicase1	4.20
	417531	NM_003157	Hs.1087	serine/threonine kinase 2	pkinase	4.12
50	428695	AI355647	Hs.189999	purinergic receptor (family A group 5)	7tm_1	3.91
	410011	AB020641	Hs.57856	PFTAIRe protein kinase 1	pkinase	3.91
	424850	AA151057	Hs.153498	chromosome 18 open reading frame 1	ldl_recept_a	3.82
	412350	AI659306	Hs.73826	protein tyrosine phosphatase, non-recept	Y_phosphatase,Band_41,PDZ	3.70
	447397	BE247676	Hs.18442	E-1 enzyme	Hydrolase	3.68
55	452946	X95425	Hs.31092	EphA5	EPH_lbd,fn3,pkinase,SAM	3.66
	427144	X95097	Hs.2126	vasoactive intestinal peptide receptor 2	7tm_2	3.65
	443775	AF291664	Hs.204732	matrix metalloproteinase 26	Peptidase_M10	3.56
	457940	AL360159	Hs.306517	Homo sapiens TRlpartite motif protein ps	SPRY,7tm_1	3.52
	418250	U29926	Hs.83918	adenosine monophosphate deaminase (isofo	A_deaminase	3.51
60	413435	X51405	Hs.75360	carboxypeptidase E	Zn_carbOpept	3.46
	447210	AF035269	Hs.17752	phosphatidylserine-specific phospholipas	lipase	3.46

TABLE 8: 136 GENES SIGNIFICANTLY DOWN-REGULATED IN PROSTATE CANCER COMPARED TO NORMAL PROSTATE

Table 8 shows 136 genes significantly down-regulated in prostate cancer compared to normal prostate. These were selected from 59680 probesets on the Affymetrix/Eos Hu03 GeneChip array such that the ratio of "average" normal prostate to "average" prostate cancer tissues was greater than or equal to 2. The "average" normal prostate level was set to the mean amongst 4 normal prostate tissues. The "average" prostate cancer level was set to the 85th percentile amongst 73 tumor samples. In order to remove gene-specific background levels of non-specific hybridization, the 10th percentile value amongst all the tissues was subtracted from both the numerator and the denominator before the ratio was evaluated.

Pkey: Unique Eos probeset identifier number
 ExAccn: Exemplar Accession number, Genbank accession number
 UnigenelD: Unigene number
 Unigene Title: Unigene gene title
 R1: Ratio of normal prostate to prostate cancer

Pkey	ExAccn	UnigenelD	Unigene Title	R1
425932	M81650	Hs.1968	semenogelin I	57.69
425545	N98529	Hs.158295	Human mRNA for myosin light chain 3 (MLC	19.70
426752	X69490	Hs.172004	titin	15.25
442082	R41823	Hs.7413	ESTs; calsyntenin-2	10.05
407245	X90568	Hs.172004	titin	9.38
422711	D60641	Hs.21739	Homo sapiens mRNA; cDNA DKFZp586I1518 (f	9.05
420813	X51501	Hs.99949	prolactin-induced protein	8.18
411987	AA375975	Hs.183380	*ESTs, Moderately similar to ALU7_HUMAN	7.45
404567				5.62
416030	H15261	Hs.21948	ESTs	5.51
444892	AI620617	Hs.148565	ESTs	5.27
444573	AW043590	Hs.225023	ESTs	5.20
428068	AW016437	Hs.233462	ESTs	5.08
437440	AA846804	Hs.123694	ESTs	4.95
404113				4.75
452279	AA286844	Hs.61260	hypothetical protein FLJ13164	4.75
421058	AW297967	Hs.188181	ESTs	4.63
445592	AV654382	Hs.17947	*ESTs, Weakly similar to K02F3.10 [C.ele	4.53
405163				4.49
405227				4.45
454059	NM_003154	Hs.37048	statherin	4.45
450152	AI138635	Hs.22968	ESTs	4.40
407013	U35637		*gb:Human nebulin mRNA, partial cds"	4.03
403612				4.02
440089	AA864468	Hs.135646	ESTs	4.00
408988	AL119844	Hs.49476	Homo sapiens clone TUA8 Cri-du-chat regi	3.98
436726	AA324975	Hs.128993	*ESTs, Weakly similar to KIAA0465 protei	3.95
459367	BE148877		*gb:CM4-HT0244-111199-040-h12 HT0244 Hom	3.95
427318	AF186081	Hs.175783	zinc transporter	3.92
411762	AW860972		*gb:QV0-CT0387-180300-167-h07 CT0387 Hom	3.85
418668	AW407987	Hs.87150	Human clone A9A2BR11 (CAC)n/(GTG)n repea	3.75
458311	AF069478		*gb:AF069478 Homo sapiens astrocytoma li	3.61
403649				3.60
419682	H13139	Hs.92282	paired-like homeodomain transcription fa	3.58
412519	AA196241	Hs.73980	*troponin T1, skeletal, slow"	3.51
414206	AW276887	Hs.46609	ESTs	3.45
427419	NM_000200	Hs.177888	histatin 3	3.37
420777	AA280223	Hs.130865	ESTs	3.35
428134	AA421773	Hs.161008	ESTs	3.31
450218	R02018	Hs.168640	*Ank, mouse, homolog of"	3.30
433474	AI192195	Hs.147174	*EST, Highly similar to ubiquitin-protei	3.30
418833	AW974899	Hs.292776	ESTs	3.26
400440	X83957	Hs.83870	nebulin	3.16

	413778	AA090235	Hs.75535	*myosin, light polypeptide 2, regulatory	3.06
	423151	AW838068		*gb:QV3-LT0048-010300-109-f02 LT0048 Hom	3.05
	445060	AA830811	Hs.88808	ESTs	2.98
	457065	AI476318	Hs.192480	ESTs	2.95
5	432456	H00093		*gb:ph8f12u_19/iTV Outward Alu-primed hn	2.92
	405678				2.85
	406707	S73840	Hs.931	*myosin, heavy polypeptide 2, skeletal m	2.81
	444105	AW189097	Hs.166597	ESTs	2.78
	433968	AL157518	Hs.90421	PRO2463 protein	2.73
10	438522	AA809431	Hs.258886	ESTs	2.73
	436562	H71937	Hs.169756	*complement component 1, s subcomponent"	2.68
	412417	AA102268	Hs.42175	ESTs	2.67
	455590	BE072259		*gb:QV4-BT0536-271299-059-g04 BT0536 Hom	2.65
	415380	F07953	Hs.16085	putative G-protein coupled receptor	2.65
15	428729	AL162331	Hs.191436	hypothetical protein FLJ10619	2.64
	408537	AW207734		*gb:Ul-H-BI2-age-h-01-0-UI.s1 NCI_CGAP_S	2.63
	424706	AA741336	Hs.152108	transcriptional unit N143	2.63
	413212	BE072092		*gb:PM4-BT0532-160200-003-b11 BT0532 Hom	2.63
20	406704	M21665	Hs.929	*myosin, heavy polypeptide 7, cardiac mu	2.62
	437507	AA758538	Hs.246882	ESTs	2.60
	410384	AI933794	Hs.42745	ESTs	2.58
	408074	R20723	Hs.124764	ESTs	2.58
	436653	AA829828	Hs.292402	ESTs	2.52
25	458090	AI282149	Hs.56213	*ESTs, Highly similar to FXD3_HUMAN FORK	2.51
	432003	AI689154	Hs.122972	ESTs	2.50
	436915	AA737400	Hs.142230	ESTs	2.50
	410028	AW576454	Hs.258553	ESTs	2.46
	448920	AW408009	Hs.22580	alkylglycerone phosphate synthase	2.45
30	422046	AI638562		*gb:ts50a10.x1 NCI_CGAP_Ut1 Homo sapiens	2.44
	451122	AA015767	Hs.193587	ESTs	2.40
	422646	H87863	Hs.151380	ESTs	2.36
	451237	AW600293		*gb:EST00049 pGEM-T library Homo sapiens	2.36
	400001			AFFX control: BioB-3	2.36
35	415835	Z45365		*gb:HSC2NF061 normalized infant brain cD	2.36
	439706	AW872527	Hs.59761	ESTs	2.36
	423341	AW242394	Hs.252495	ESTs	2.36
	436486	AA742221	Hs.120633	ESTs	2.35
	407449	AJ002784		gb:Homo sapiens mRNA; fetal brain cDNA 5	2.33
40	430573	AA744550	Hs.136345	ESTs	2.32
	401974				2.31
	443356	AL044498	Hs.133262	*ESTs, Weakly similar to PH0217 reverse	2.31
	430751	NM_012471	Hs.247868	transient receptor potential channel 5	2.25
	439128	AI949371	Hs.153089	ESTs	2.25
45	448765	R15337	Hs.21958	*Homo sapiens cDNA FLJ10532 fis, clone N	2.25
	451130	AI762250	Hs.211347	ESTs	2.24
	405420				2.23
	455029	AW851258		*gb:IL3-CT0220-160200-066-H06 CT0220 Hom	2.23
	438224	AA933999		*gb:on91f04.s1 Soares_NFL_T_GBC_S1 Homo	2.23
50	407764	BE008347		*gb:CM0-BN0154-080400-325-h04 BN0154 Hom	2.23
	413549	BE252470		*gb:601108292F1 NIH_MGC_16 Homo sapiens	2.23
	437010	AA741368	Hs.291434	ESTs	2.23
	435111	AI914279	Hs.213740	ESTs	2.22
	403375				2.21
	455060	AW853441		*gb:RC1-CT0252-030100-023-g09 CT0252 Hom	2.21
55	409792	AW854153		*gb:RC3-CT0254-060400-029-d03 CT0254 Hom	2.20
	421154	AA284333	Hs.287631	*Homo sapiens cDNA FLJ14269 fis, clone P	2.19
	401963				2.18
	435034	AF168711	Hs.159397	x 010 protein	2.18
	448996	AW998989	Hs.105749	KIAA0553 protein	2.18
60	436816	AW297599	Hs.255667	ESTs	2.17
	442252	AI733395	Hs.129124	ESTs	2.17
	419310	AA236233	Hs.188716	ESTs	2.16
	418579	H91800	Hs.124156	ESTs	2.16
	423315	R54109	Hs.26096	ESTs	2.16
65	432744	AA988835	Hs.38664	ESTs	2.15
	424492	AI133482	Hs.165210	ESTs	2.15
	424770	AA425562		*gb:zw46a05.r1 Soares_total_fetus_Nb2HF8	2.15
	437101	AA744518	Hs.120610	ESTs	2.15
	428793	AC004957	Hs.298975	*ESTs, Highly similar to collapsin-2-lik	2.15

	415708	H56475		"gb:yt87d11.r1 Soares_pineal_gland_N3HPG	2.13
	459619				2.12
	427506	AK000134	Hs.179100	hypothetical protein FLJ20127	2.12
	452508	AA804174	Hs.184354	ESTs	2.10
5	410881	AW809157		"gb:RC0-ST0118-041099-031-c07_1 ST0118 Homo sapiens cDNA, mRNA sequence"	2.10
	403087				2.10
	403869				2.10
	445028	D81194	Hs.282499	ESTs	2.10
10	447884	H29505		"gb:ym60d10.r1 Soares infant brain 1NIB Homo sapiens cDNA clone 5', mRNA sequence"	2.10
	414575	H11257	Hs.295233	ESTs	2.09
	420351	BE218221	Hs.190044	ESTs	2.08
	426998	BE274360		"gb:601121068F1 NIH_MGC_20 Homo sapiens cDNA clone 5', mRNA sequence"	2.08
	405455				2.08
15	423843	AA332652		"gb:EST36627 Embryo, 8 week I Homo sapiens cDNA 5' end similar to similar to monoamine oxidase B, mRNA sequence"	2.08
	406135				2.07
	427046	BE246180	Hs.121385	ESTs	2.07
	403493				2.05
20	444514	AI682905	Hs.270431	"ESTs, Weakly similar to ALU1_HUMAN ALU SUBFAMILY J SEQUENCE CONTAMINATION WARNING ENTRY [H.sapiens]"	2.05
	435884	AA701443	Hs.192868	ESTs	2.05
	419629	AB020695	Hs.91662	KIAA0888 protein	2.03
	405900				2.03
	457350	AW974438	Hs.194136	"ESTs, Moderately similar to AF091457 1 zinc finger protein RIN ZF [R.norvegicus]"	2.02
25	400007			AFFX control: BioDn-5	2.01
	406978	M64358		"gb:Human rhom-3 gene, exon."	2.00

TABLE 8A shows the accession numbers for those primekeys lacking a unigeneID in Table 8. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

10	Pkey:	Unique Eos probeset identifier number	
	CAT number:	Gene cluster number	
	Accession:	Genbank accession numbers	
<hr/>			
15	Pkey	CAT number	Accessions
	407764	1014849_1	BE008347 BE008320 BE083307 BE083311 AW075968
	408537	1064753_1	AW207734 D60164 D81150 D81078 D61356 AW996804
	409792	1154677_1	AW854153 AW500210 BE145772 AW501310
20	410881	1225682_1	AW809157 AW812181 AW812175 AW812172 AW812161 AW812165
	411762	1256906_1	AW860972 AW862598 AW862599 AW860988 AW860983 AW860898 AW860925 AW860922 AW860986 AW860984 AW860989
	413212	1353792_1	BE072092 BE072106 BE072086 BE072098 BE072103
	413549	1375933_2	BE252470 BE147573
	415708	1548209_1	H56475 F29401 F34552
25	415835	1558511_1	Z45365 R25905 H05203 T77496
	422046	210744_1	AI638562 T16929 H13401 F07773 R55836
	423151	225415_1	AW838068 AW837986 AW838067 AA322487 AW837936
	423843	232510_1	AA332652 AA331633 AW999369 AW902993 BE170475 AA378845 AW964175 AI475221
	424770	243504_1	AA425562 AI880208 AA346646 N22655 AW811775 AW811786
30	426998	274259_1	BE274360
	432456	347718_2	H00093 H00079 H00070 H00054 H00049 H00063 AW905306 AW905241 AW905410 AW905307 AW905411 AW905240
	AW905210		AW905352 AW905304 AW905239 AW905242 AW905243 H00087
	438224	452656_1	AA933999 AA781181
35	447884	740749_1	H29505 R18575 Z43580 T48738 AI435454 BE004683
	451237	863269_1	AW600293 AI767468
	455029	1249374_1	AW851258 AW851435 AW851106 AW851421
	455060	1251259_1	AW853441 BE145228 BE145218 BE145162 BE145283
	455590	1335127_1	BE072259 BE072230 BE007911
40	458311	543550_1	AF069478 AF069479 AF069480

TABLE 8B shows the genomic positioning for those primekeys lacking unigene ID's and accession numbers in table 8. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

5

Pkey: Unique number corresponding to an Eos probeset
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
 Strand: Indicates DNA strand from which exons were predicted.
 Nt_position: Indicates nucleotide positions of predicted exons.

10

15

Pkey	Ref	Strand	Nt_position
401963	3126783	Plus	51382-51521
401974	3126777	Plus	85330-85683
403087	8954241	Plus	169511-169795
403375	9255944	Minus	92554-92795
403493	7341425	Plus	157568-159084
403612	8469060	Minus	94723-94859
403649	8705159	Minus	27141-27247
403869	7280046	Minus	34379-34583
404113	9588571	Minus	13446-13646
404567	7249169	Minus	101320-101501
405163	9966267	Minus	161171-161299
405227	6731245	Minus	22550-22802
405420	7211837	Minus	13428-13582
405455	7656675	Plus	134112-134671
405678	4079670	Plus	151821-152027
405900	6758795	Minus	71181-71535
406135	9164918	Minus	65489-65715

20

25

30

TABLE 9: 1001 GENES SIGNIFICANTLY UP-REGULATED IN NORMAL PROSTATE COMPATED TO PROSTATE CANCER

Table 9 shows 1001 genes significantly up-regulated in prostate cancer compared to normal prostate. These were selected from 59680 probesets on the Affymetrix/Eos Hu03 GeneChip array such that the ratio of "average" normal prostate to "average" prostate cancer tissues was greater than or equal to 8.14. The "average" normal prostate level was set to the mean amongst 4 normal prostate tissues. The "average" prostate cancer level was set to the 85th percentile amongst 73 tumor samples. In order to remove gene-specific background levels of non-specific hybridization, the 10th percentile value amongst all the tissues was subtracted from both the numerator and the denominator before the ratio was evaluated.

Pkey: Unique Eos probeset identifier number
 ExAccn: Exemplar Accession number, Genbank accession number
 UnigenelD: Unigene number
 Unigene Title: Unigene gene title
 R1: Ratio of prostate cancer to normal prostate

Pkey	ExAccn	UnigenelD	Unigene Title	R1
451002	AA013299	Hs.8018	ESTs, Weakly similar to ALU3_HUMAN ALU S	1684.00
435596	AA689465	Hs.188999	ESTs	738.00
443576	AI078027	Hs.169338	ESTs	246.86
434247	AA928116	Hs.272065	ESTs	245.20
400452	AK000185		gb:Homo sapiens cDNA FLJ20178 fis, clone	222.00
405932				221.33
427906	AA864330	Hs.166520	ESTs	212.00
443685	AI686550	Hs.174481	ESTs	163.20
451554	AI474866	Hs.193237	ESTs	149.45
418323	NM_002118	Hs.1162	major histocompatibility complex, class	126.11
429480	M36860	Hs.9295	elastin (supraaortic aortic stenosis,	123.27
426025	AW138330	Hs.233778	ESTs	120.00
418917	X02994	Hs.1217	adenosine deaminase	106.75
404407				105.71
442027	AI652926	Hs.128395	ESTs	100.53
433704	AA608684	Hs.121705	ESTs, Moderately similar to ALUC_HUMAN I	94.00
453758	U83527		gb:HSU83527 Human fetal brain (M.Lovett)	89.18
415354	F06495		gb:HSC1AB051 normalized infant brain cDN	87.73
424239	M67439	Hs.143526	dopamine receptor D5	86.82
444143	AW747996	Hs.160999	ESTs	86.43
401672				77.26
430590	AW383947	Hs.246381	CD68 antigen	68.47
411972	BE074959		gb:PM0-BT0582-310100-001-f08 BT0582 Homo	68.00
448992	AI766053	Hs.188346	ESTs	61.26
408828	BE540279		gb:601059857F1 NIH_MGC_10 Homo sapiens c	57.71
409653	AW451693	Hs.220826	ESTs	56.40
402964				54.67
422673	N59027		gb:yv59d11.r1 Soares fetal liver spleen	54.00
422568	AA372275	Hs.279800	Homo sapiens cDNA FLJ11383 fis, clone HE	54.00
438907	R32704	Hs.301298	ESTs	52.96
405172				52.96
444897	AW137088	Hs.144857	ESTs	52.32
458019	AW592931	Hs.256298	ESTs	51.63
405275	AB028989	Hs.88500	mitogen-activated protein kinase 8 inter	50.98
457815	AA703679	Hs.106999	ESTs, Weakly similar to SYT5_HUMAN SYNAP	49.60
424385	AA339666		gb:EST44776 Fetal brain I Homo sapiens c	48.90
407172	T54095		gb:ya92c05.s1 Stratagene placenta (93722	47.98
428202	AA424163	Hs.156895	ESTs	46.83
435672	AI700148	Hs.283626	ESTs	43.57
420283	AA485224	Hs.57734	G protein-coupled receptor kinase-intera	43.00
417016	AA837098	Hs.269933	ESTs	42.70
438854	AF074994	Hs.24240	ESTs	42.67

	406134			42.43
	457319	AA480895	Hs.201552 ESTs, Weakly similar to T17288 hypotheti	42.31
	409314	AA070266	gb:zm69d04.r1 Stratagene neuroepithelium	42.25
	401124			41.61
5	429316	AI371157	Hs.178538 ESTs	40.00
	420317	AB006628	Hs.96485 KIAA0290 protein	39.64
	457586	AW062439	gb:MRO-CT0060-120899-001-f08 CT0060 Homo	39.60
	417407	AA923278	Hs.290905 ESTs, Weakly similar to protease [H.sapi	38.73
	430269	BE221682	Hs.178364 ESTs	38.06
10	439602	W79114	Hs.58558 ESTs	36.69
	433686	AA604799	Hs.136528 ESTs, Moderately similar to ALU1_HUMAN A	36.29
	417993	AW963705	Hs.295806 ESTs, Weakly similar to ALU7_HUMAN ALU S	36.18
	428214	AA936282	Hs.120397 ESTs	36.10
	416908	AA333990	Hs.80424 coagulation factor XIII, A1 polypeptide	36.08
15	426264	BE314852	Hs.168694 hypothetical protein FLJ10257	36.00
	415911	H08796	Hs.124952 ESTs	36.00
	457502	AA076049	Hs.274415 Homo sapiens cDNA FLJ10229 fis, clone HE	35.23
	421566	NM_000399	Hs.1395 early growth response 2 (Krox-20 (Drosop	35.20
	401468			34.89
20	458561	AI220150	Hs.211195 ESTs	34.60
	433601	BE350738	Hs.123993 ESTs, Weakly similar to T00366 hypotheti	33.24
	454977	AW848032	gb:IL3-CT0214-231299-053-D11 CT0214 Homo	32.96
	402828			32.93
	414522	AW518944	Hs.76325 Homo sapiens cDNA: FLJ23125 fis, clone L	31.76
25	402842			31.68
	421245	AA285363	gb:HTH280 HTC DL1 Homo sapiens cDNA 5'/3'	31.59
	401631	F05183	Hs.1799 CD1D antigen, d polypeptide	31.26
	408057	AW139565	gb:UL-H-BI1-aea-d-04-0-UI.s1 NCI_CGAP_Su	31.24
	408069	H81795	gb:ys68a10.r1 Soares retina N2b4HR Homo	31.20
30	438694	T87479	Hs.291797 ESTs	31.09
	449156	AF103907	Hs.171353 prostate cancer antigen 3	29.78
	428796	AU076734	Hs.193665 solute carrier family 28 (sodium-coupled	29.76
	452549	AI907039	gb:PM-BT134-020499-566 BT134 Homo sapien	29.59
	410129	BE244074	Hs.285531 regulator of Fas-induced apoptosis	29.53
35	414464	AI870175	Hs.13957 ESTs	29.47
	412326	R07566	Hs.73817 Small inducible cytokine A3 (homologous	29.22
	459081	W07808	gb:zb03a12.r1 Soares_fetal_lung_NbHL19W	29.20
	448702	AW102670	Hs.122464 ESTs	29.13
40	451939	U80456	Hs.27311 single-minded (Drosophila) homolog 2	28.74
	443412	W84893	Hs.9305 angiotensin receptor-like 1	28.61
	457324	AB028990	Hs.243901 KIAA1067 protein	28.24
	424247	X14008	Hs.234734 lysozyme (renal amyloidosis)	28.18
	457140	AI279960	Hs.178140 ESTs	28.12
	444151	AW972917	Hs.128749 alpha-methylacyl-CoA racemase	28.06
45	457669	AW104257	Hs.123426 ESTs, Weakly similar to putative serine/	27.61
	412429	AV650262	Hs.75765 GRO2 oncogene	27.36
	405495			27.33
	406516			27.25
	407997	AW135429	Hs.243577 ESTs	26.96
50	442115	AW452332	Hs.257554 ESTs	26.36
	409038	T97490	Hs.50002 small inducible cytokine subfamily A (Cy	26.34
	402838			26.32
	449846	AI979284	Hs.200552 ESTs	26.21
	417153	X57010	Hs.81343 collagen, type II, alpha 1 (primary oste	26.20
55	439792	NM_014856	Hs.6684 KIAA0476 gene product	25.91
	450096	AI682088	Hs.223368 ESTs	25.60
	424196	AL133660	Hs.142926 Homo sapiens mRNA; cDNA DKFZp434M0927 (f	25.57
	414246	BE391090	Hs.280278 EST	25.57
	420848	NM_005188	Hs.99980 Cas-Br-M (murine) ecotropic retroviral t	25.48
60	424778	AA251048	Hs.153042 lymphocyte antigen 9	25.42
	409126	AA063426	gb:zf70c08.s1 Soares_pineal_gland_N3HPG	25.25
	443936	AW083491	Hs.31196 ESTs	25.22
	419392	W28573	gb:51f10 Human retina cDNA randomly prim	25.01
	411201	T74588	Hs.8509 ESTs, Weakly similar to CO3_HUMAN COMPLE	24.85
65	422940	BE077458	gb:RC1-BT0606-090500-015-b04 BT0606 Homo	24.76
	437571	AA760894	Hs.153023 ESTs	24.74
	433973	AI014723	Hs.131770 ESTs	24.57
	422416	BE019557	Hs.11900 Human DNA sequence from clone RP4-583P15	24.53
	421552	AF026692	Hs.105700 secreted frizzled-related protein 4	24.49

	443668	U25758	Hs.134584	ESTs	24.49
	424800	AL035588	Hs.153203	MyoD family inhibitor	24.10
	453633	AA357001	Hs.34045	hypothetical protein FLJ20764	24.04
	430565	AL122081	Hs.244343	cadherin related 23	24.00
5	433694	AI208611	Hs.12066	Homo sapiens cDNA FLJ11720 fis, clone HE	23.89
	451045	AA215672		gb:zr96e09.s1 NCI_CGAP_GCB1 Homo sapiens	23.83
	408583	AW449674	Hs.47359	ESTs	23.73
	444040	AF204231	Hs.182982	golgin-67	23.62
	414182	AA136301		gb:zk93g04.s1 Soares_pregnant_uterus_NbH	23.39
10	418678	NM_001327	Hs.167379	cancer/testis antigen	23.20
	408380	AF123050	Hs.44532	diubiquitin	22.68
	456076	BE243877	Hs.76941	ATPase, Na+/K+ transporting, beta 3 poly	22.65
	418299	AA279530	Hs.83968	integrin, beta 2 (antigen CD18 (p95), ly	22.38
	444917	R68651	Hs.144997	ESTs	22.26
15	444381	BE387335	Hs.283713	ESTs	22.08
	415788	AW628686	Hs.78851	KIAA0217 protein	22.04
	410896	AW809637		gb:MR4-ST0124-261099-015-b07 ST0124 Homo	22.00
	412978	AI431708	Hs.820	homeo box C6	21.95
	458418	AV653846	Hs.126261	Homo sapiens Chromosome 16 BAC clone CIT	21.94
20	454791	BE071874		gb:RC2-BT0522-120200-014-a06 BT0522 Homo	21.84
	408748	J05500	Hs.47431	spectrin, beta, erythrocytic (includes s	21.26
	416011	H14487		gb:ym18c10.r1 Soares infant brain 1NIB H	21.24
	440474	AI207936	Hs.7195	gamma-aminobutyric acid (GABA) A recepto	21.14
	447047	AI623698	Hs.246306	Homo sapiens cDNA: FLJ23529 fis, clone L	21.11
25	426793	X89887	Hs.172350	HIR (histone cell cycle regulation defec	21.10
	409841	AW502139		gb:UL-HF-BR0p-ajr-e-05-0-UL.r1 NIH_MGC_5	21.07
	405685				20.90
	457359	AI983207	Hs.192481	ESTs, Weakly similar to SYPH_HUMAN SYNAP	20.84
	423067	AA321355	Hs.285401	ESTs	20.74
30	422355	AW403724	Hs.140	immunoglobulin heavy constant gamma 3 (G	20.73
	401201				20.73
	458278	W28912	Hs.129019	ESTs	20.68
	439097	H66948		gb:yr86d10.r1 Soares fetal liver spleen	20.67
	414875	H42679	Hs.77522	major histocompatibility complex, class	20.66
35	400926				20.66
	451355	NM_004197	Hs.444	serine/threonine kinase 19	20.64
	446982	AW500221	Hs.43616	Homo sapiens mRNA for FLJ00029 protein,	20.61
	417105	X60992	Hs.81226	CD6 antigen	20.61
	405777				20.51
40	424123	AW966158	Hs.58582	Homo sapiens cDNA FLJ12702 fis, clone NT	20.20
	425009	X58288	Hs.154151	protein tyrosine phosphatase, receptor t	20.10
	443271	BE568568	Hs.195704	ESTs	19.98
	421064	AI245432	Hs.101382	tumor necrosis factor, alpha-induced pro	19.98
	418819	AA228776	Hs.191721	ESTs	19.94
45	457595	AA584854		gb:no09h11.s1 NCI_CGAP_Phe1 Homo sapiens	19.90
	404426				19.84
	412571	U43143	Hs.74049	fms-related tyrosine kinase 4	19.79
	431457	NM_012211	Hs.256297	integrin, alpha 11	19.62
	414002	NM_006732	Hs.75678	FBJ murine osteosarcoma viral oncogene h	19.57
50	418994	AA296520	Hs.89546	Selectin E (endothelial adhesion molecu	19.56
	437158	AW090198	Hs.4779	KIAA1150 protein	19.52
	437866	AA156781	Hs.83992	ESTs	19.44
	417421	AL138201	Hs.82120	nuclear receptor subfamily 4, group A, m	19.34
	433057	X15675	Hs.296832	Human pTR7 mRNA for repetitive sequence	19.22
55	421730	AW449808	Hs.164036	glucosamine (N-acetyl)-6-sulfatase (Sanf	19.21
	456557	AA284477	Hs.96618	ESTs	18.77
	440806	AI247422	Hs.129966	ESTs	18.76
	439845	AL355743	Hs.56663	Homo sapiens EST from clone 41214, full	18.65
	416155	AI807264	Hs.205442	ESTs, Weakly similar to AF117610 1 inner	18.64
60	437820	AA769062	Hs.16029	ESTs, Weakly similar to alternatively sp	18.62
	450923	AW043951	Hs.38449	ESTs	18.59
	418329	AW247430	Hs.84152	cystathionine-beta-synthase	18.58
	424537	AI673027	Hs.143271	ESTs	18.55
	447742	AF113925	Hs.19405	caspase recruitment domain 4	18.52
65	415251	R42863	Hs.7124	ESTs	18.47
	440770	AA912815	Hs.222078	ESTs	18.40
	407711	AI085846	Hs.25522	ESTs	18.32
	427157	U51166	Hs.173824	thymine-DNA glycosylase	18.28
	409847	AW501751	Hs.279733	ESTs	18.15

	417240	N57568	Hs.176028	EST	18.13
	435732	AF229178	Hs.123136	leucine rich repeat and death domain con	18.12
	436896	AW977385	Hs.278615	ESTs	18.12
	432485	N90866	Hs.276770	CDW52 antigen (CAMPATH-1 antigen)	17.90
5	429490	AI971131	Hs.293684	ESTs, Weakly similar to alternatively sp	17.82
	429984	AL050102	Hs.227209	DKFZP586F1019 protein	17.82
	449214	AI889114	Hs.195663	ESTs	17.75
	433867	AK000596	Hs.3618	hippocalcin-like 1	17.72
10	431735	AW977724	Hs.75968	thymosin, beta 4, X chromosome	17.71
	401515				17.67
	444045	AI097439	Hs.135548	ESTs	17.58
	442754	AL045825	Hs.210197	ESTs	17.55
	426559	AB001914	Hs.170414	paired basic amino acid cleaving system	17.54
15	432415	T16971	Hs.289014	ESTs	17.50
	427829	AI188225	Hs.127462	ESTs	17.50
	432516	R08003	Hs.188013	ESTs	17.44
	435259	AA152106	Hs.4859	cyclin L ania-6a	17.36
	414989	T81668		gb:yd29c04.r1 Soares fetal liver spleen	17.31
20	444880	AW118683	Hs.154150	ESTs	17.30
	417651	R06874	Hs.268628	ESTs	17.27
	453457	AL037103	Hs.270599	ESTs, Weakly similar to unnamed protein	17.22
	424246	AW452533	Hs.143604	Kaiso	17.22
	419078	M93119	Hs.89584	insulinoma-associated 1	17.18
25	417696	BE241624	Hs.82401	CD69 antigen (p60, early T-cell activati	17.14
	431117	AF003522	Hs.250500	delta (Drosophila)-like 1	17.14
	455254	AW877015		gb:QV2-PT0010-250300-096-f12 PT0010 Homo	17.14
	425782	U66468	Hs.159525	cell growth regulatory with EF-hand doma	17.12
	426678	H08170	Hs.113755	ESTs	17.12
30	426403	NM_000361	Hs.2030	thrombomodulin	17.01
	425905	AB032959	Hs.161700	KIAA1133 protein	17.00
	438867	AW451157	Hs.181157	ESTs	16.98
	420940	AA830664	Hs.143974	ESTs	16.94
	459234	AI940425		gb:CM0-CT0052-150799-024-c04 CT0052 Homo	16.92
35	404756				16.91
	422247	U18244	Hs.113602	solute carrier family 1 (high affinity a	16.90
	420568	F09247	Hs.167399	protocadherin alpha 5	16.88
	443559	AI076765	Hs.269899	ESTs	16.80
	438703	AI803373	Hs.31599	ESTs	16.78
40	411424	AW845985		gb:RC2-CT0163-200999-002-H08 CT0163 Homo	16.70
	402895				16.69
	422538	NM_006441	Hs.118131	5,10-methenyltetrahydrofolate synthetase	16.68
	447108	AW449602	Hs.217953	ESTs, Moderately similar to NK-TUMOR REC	16.65
	448520	AB002367	Hs.21355	doublecortin and CaM kinase-like 1	16.54
45	438567	AW451955	Hs.153065	ESTs	16.52
	407811	AW190902	Hs.40098	cysteine knot superfamily 1, BMP antagon	16.50
	410721	R23534	Hs.2730	heterogeneous nuclear ribonucleoprotein	16.50
	437133	AB018319	Hs.5460	KIAA0776 protein	16.40
	408182	AA047854		gb:zf49g04.r1 Soares retina N2b4HR Homo	16.32
50	417315	AI080042	Hs.180450	ribosomal protein S24	16.30
	431840	AA534908	Hs.2860	POU domain, class 5, transcription facto	16.28
	439882	AA847856	Hs.124565	ESTs	16.20
	418277	AW135221	Hs.130812	ESTs	16.09
	410688	AW796342		gb:PM2-UM0027-230200-002-h02 UM0027 Homo	16.04
55	420120	AL049610	Hs.95243	transcription elongation factor A (SII)-	16.04
	429597	NM_003816	Hs.2442	a disintegrin and metalloproteinase doma	16.02
	447033	AI357412	Hs.157601	EST - not in UniGene	16.02
	421684	BE281591	Hs.106768	hypothetical protein FLJ10511	15.94
	408599	AA055800	Hs.222933	ESTs	15.93
60	446012	AV656098	Hs.172382	hypothetical protein FLJ20001	15.86
	409671	AA076769		gb:7B02B10 Chromosome 7 Fetal Brain cDNA	15.85
	405934				15.84
	426108	AA622037	Hs.166468	programmed cell death 5	15.84
	416208	AW291168	Hs.41295	ESTs	15.48
	410708	AA534370	Hs.154088	Homo sapiens cDNA: FLJ22756 fis, clone K	15.42
65	447342	AI199268	Hs.19322	ESTs; Weakly similar to !!!! ALU SUBFAMI	15.38
	454563	AW807530		gb:CM0-ST0081-130999-054-d02 ST0081 Homo	15.37
	411507	AW850140		gb:IL3-CT0219-261099-023-D11 CT0219 Homo	15.36
	438170	AI916685	Hs.194601	ESTs	15.29
	416292	AA179233	Hs.42390	nasopharyngeal carcinoma susceptibility	15.26

	406638	M13861	gb:Human T-cell receptor active beta-cha	15.26
	446686	AW138043	Hs.156307 ESTs	15.25
	434485	AI623511	Hs.118567 ESTs	15.24
5	441188	AW292830	Hs.255609 ESTs	15.22
	444172	BE147740	Hs.104558 ESTs	15.22
	409521	BE244854	Hs.159578 Homo sapiens mRNA for FLJ00020 protein,	15.16
	420748	AA279956	Hs.88672 ESTs	15.14
	422583	AA410506	Hs.118578 H.sapiens mRNA for ribosomal protein L18	15.14
10	424240	AB023185	Hs.143535 calcium/calmodulin-dependent protein kin	15.12
	451118	AI862096	Hs.60640 ESTs	15.12
	437495	BE177778	gb:RC1-HT0598-310300-012-f07 HT0598 Homo	15.12
	445467	AI239832	Hs.15617 ESTs, Weakly similar to ALU4_HUMAN ALU S	15.06
	418305	AW006783	Hs.6686 ESTs	15.03
	402812			15.02
15	436851	AA732480	Hs.293581 ESTs	15.00
	400991			15.00
	415752	BE314524	Hs.78776 Human putative transmembrane protein (nm	14.96
	429900	AA460421	Hs.30875 ESTs	14.90
	403683			14.84
20	430315	NM_004293	Hs.239147 guanine deaminase	14.80
	451952	AL120173	Hs.301663 ESTs	14.72
	424687	J05070	Hs.151738 matrix metalloproteinase 9 (gelatinase B	14.69
	447229	BE617135	gb:601441677F1 NIH_MGC_65 Homo sapiens c	14.67
	425818	AB021225	Hs.159581 matrix metalloproteinase 17 (membrane-in	14.65
25	448553	AI638449	Hs.173031 ESTs	14.63
	431089	BE041395	Hs.283676 ESTs, Weakly similar to unknown protein	14.60
	459145	AI903354	gb:RC-BT029-100199-117 BT029 Homo sapien	14.55
	449650	AF055575	Hs.297647 ESTs, Moderately similar to calcium chan	14.54
	400952			14.46
30	445885	AI734009	Hs.127699 EST cluster (not in UniGene)	14.44
	407938	AA905097	Hs.85050 phospholamban	14.42
	431676	AI685464	Hs.292638 ESTs	14.40
	437210	AA311443	Hs.293563 Homo sapiens mRNA; cDNA DKFZp586E2317 (f	14.36
	451900	AB023199	Hs.27207 KIAA0982 protein	14.36
35	445800	AA126419	Hs.301632 ESTs	14.32
	412368	AW945992	Hs.181125 immunoglobulin lambda locus	14.31
	409055	AW304028	Hs.300578 ESTs	14.23
	408763	W57550	Hs.301526 Homo sapiens cDNA FLJ13181 fis, clone NT	14.22
	446734	AL049278	Hs.16074 Homo sapiens mRNA; cDNA DKFZp564I153 (fr	14.22
40	413551	BE242639	Hs.75425 ubiquitin associated protein	14.22
	421913	AI934365	Hs.109439 osteoglycin (osteoinductive factor, mime	14.22
	452712	AW838616	gb:RC5-LT0054-140200-013-D01 LT0054 Homo	14.22
	451468	AW503398	Hs.210047 ESTs	14.16
	406038	Y14443	Hs.88219 zinc finger protein 200	14.14
45	424909	S78187	Hs.153752 cell division cycle 25B	14.07
	434078	AW880709	Hs.283683 EST	14.07
	415254	AI815831	Hs.184378 ESTs	14.05
	418196	AI745649	Hs.26549 ESTs, Weakly similar to T00066 hypotheti	14.02
	410020	T86315	Hs.728 ribonuclease, RNase A family, 2 (liver,	13.98
50	411352	NM_002890	Hs.758 RAS p21 protein activator (GTPase activa	13.98
	429848	AF145439	Hs.225946 chemokine (C-C motif) receptor 9	13.95
	413729	BE159999	gb:QV1-HT0412-270300-123-d10 HT0412 Homo	13.90
	400125			13.88
	420319	AW406289	Hs.96593 hypothetical protein	13.85
55	448272	AI479094	Hs.170786 ESTs	13.80
	422695	AA315158	gb:EST186956 HCC cell line (matatasis t	13.80
	424565	AW102723	Hs.75295 guanylate cyclase 1, soluble, alpha 3	13.78
	458048	H30340	Hs.173705 Homo sapiens cDNA: FLJ22050 fis, clone H	13.78
	408894	AI935400	Hs.217286 ESTs	13.76
60	454093	AW860158	gb:RC0-CT0379-290100-032-b04 CT0379 Homo	13.75
	410889	X91662	Hs.66744 twist (Drosophila) homolog (acrocephalos	13.74
	457751	AI908236	gb:IL-BT166-180399-010 BT166 Homo sapien	13.72
	455131	AW857913	gb:RC0-CT0323-231199-031-b05 CT0323 Homo	13.69
	408364	AW015238	Hs.128453 ESTs	13.67
65	425907	AA365752	Hs.155965 ESTs	13.62
	402359			13.60
	401044			13.53
	409877	AW502498	Hs.157150 ESTs, Weakly similar to zinc finger prot	13.53
	423690	AA329648	Hs.23804 ESTs	13.49

	430685	AI690234	Hs.191666	ESTs, Weakly similar to reverse transcri	13.47
	414052	AW578849	Hs.283552	ESTs, Weakly similar to unnamed protein	13.46
	447858	AW080339	Hs.211911	ESTs	13.44
	435716	AI573283	Hs.38458	ESTs	13.44
5	439120	H56389		gb:yt87c03.r1 Soares_pineal_gland_N3HPG	13.43
	402788				13.40
	451591	AA886446	Hs.146278	ESTs	13.40
	405411				13.38
10	426558	AW188574	Hs.24218	ESTs	13.34
	453506	AA132818	Hs.110407	ESTs, Weakly similar to coded for by C.	13.33
	416445	AL043004	Hs.300678	Human serine/threonine kinase mRNA, part	13.32
	457084	AI074149	Hs.150905	ESTs, Weakly similar to chondroitin 4-su	13.32
	403838				13.32
	427337	Z46223	Hs.176663	Fc fragment of IgG, low affinity IIIb, r	13.30
15	434318	AW207552	Hs.116328	ESTs, Weakly similar to dJ134E15.1 [H.sa	13.28
	435193	N41359	Hs.218107	ESTs	13.28
	414756	AW451101	Hs.159489	ESTs, Moderately similar to hexokinase I	13.27
	420626	AF043722	Hs.99491	RAS guanyl releasing protein 2 (calcium	13.26
	420052	AA418850	Hs.44410	ESTs	13.25
20	414020	NM_002984	Hs.75703	small inducible cytokine A4 (homologous	13.25
	403851				13.24
	422647	W07492	Hs.157101	ESTs	13.21
	433598	AI762836	Hs.271433	ESTs, Moderately similar to ALU2_HUMAN A	13.21
25	409065	AB033113	Hs.50187	KIAA1287 protein	13.20
	435063	R21966	Hs.57734	G protein-coupled receptor kinase-intera	13.19
	439367	BE386844	Hs.248746	ESTs	13.17
	451957	AI796320	Hs.10299	Homo sapiens cDNA FLJ13545 fis, clone PL	13.16
	420569	AA278362	Hs.289062	Homo sapiens cDNA FLJ12334 fis, clone MA	13.14
	447883	BE262802	Hs.4909	dickkopf (Xenopus laevis) homolog 3	13.07
30	426490	NM_001621	Hs.170087	aryl hydrocarbon receptor	13.06
	414789	AA155859	Hs.79708	ESTs	13.05
	451418	BE387790	Hs.26369	ESTs	13.04
	443494	T99719	Hs.270404	Homo sapiens cDNA: FLJ22389 fis, clone H	13.03
	425878	AW964806	Hs.38085	ESTs, Weakly similar to putative glycine	13.02
35	431912	AI660552	Hs.154903	ESTs, Weakly similar to A56154 Abl subst	13.00
	407122	H20276	Hs.31742	ESTs	13.00
	456491	AL137466	Hs.97277	Homo sapiens mRNA; cDNA DKFZp434H1322 (f	12.99
	448172	N75276	Hs.135904	ESTs	12.98
	452144	AA032197	Hs.102558	ESTs	12.96
40	419953	BE267154	Hs.125752	ESTs	12.96
	416182	NM_004354	Hs.79069	cyclin G2	12.94
	451154	AA015879	Hs.33536	ESTs	12.93
	412257	AW903830		gb:CM4-NN1037-250400-155-h04 NN1037 Homo	12.93
	449784	AW161319	Hs.12915	ESTs	12.92
45	432695	D63480	Hs.278634	KIAA0146 protein	12.92
	454105	NM_001259	Hs.38481	cyclin-dependent kinase 6	12.92
	439093	AA534163	Hs.5476	serine protease inhibitor, Kazal type, 5	12.90
	416098	H41324	Hs.31581	ESTs, Moderately similar to ST1B_HUMAN S	12.88
	424897	D63216	Hs.153684	frizzled-related protein	12.88
50	414604	AU076649	Hs.76556	growth arrest and DNA-damage-inducible 3	12.88
	414664	AA587775	Hs.66295	Homo sapiens HSPC311 mRNA, partial cds	12.84
	452560	BE077084		gb:RC5-BT0603-220200-013-C07 BT0603 Homo	12.84
	413869	NM_000878	Hs.75596	interleukin 2 receptor, beta	12.80
	452359	BE167229	Hs.29206	Homo sapiens clone 24659 mRNA sequence	12.80
55	435886	BE265839	Hs.12126	hepatocellular carcinoma-associated anti	12.78
	445230	U97018	Hs.12451	echinoderm microtubule-associated protei	12.78
	412226	W26786		gb:15d7 Human retina cDNA randomly prime	12.77
	446619	AU076643	Hs.313	secreted phosphoprotein 1 (osteopontin,	12.76
	447769	AW873704	Hs.48764	ESTs	12.76
60	414478	AI306389	Hs.76240	adenylate kinase 1	12.76
	425383	D83407	Hs.156007	Down syndrome critical region gene 1-lik	12.68
	450704	H85157	Hs.40696	ESTs	12.66
	405856				12.66
65	412935	BE267045	Hs.75064	tubulin-specific chaperone c	12.65
	402802				12.62
	452588	AA889120	Hs.110637	Homeo box A10	12.62
	419978	NM_001454	Hs.93974	forkhead box J1	12.62
	403137				12.60
	430226	BE245562	Hs.2551	adrenergic, beta-2-, receptor, surface	12.57

	448076	AJ133123	Hs.20196	adenylate cyclase 9	12.56
	450462	F07097	Hs.300828	Homo sapiens mRNA full length insert cDN	12.54
	405236				12.52
5	409292	AA071051		gb:zm58e05.s1 Stratagene fibroblast (937	12.47
	421540	AA767669	Hs.10242	ESTs	12.47
	425840	AW978731	Hs.301824	ESTs	12.44
	443181	AI039201	Hs.54548	ESTs	12.42
	452436	BE077546	Hs.31447	ESTs	12.42
10	455183	AW984111		gb:RCO-HN0007-160300-011-f09 HN0007 Homo	12.40
	432887	AI926047	Hs.162859	ESTs	12.37
	410494	M36564	Hs.64016	protein S (alpha)	12.36
	439024	R96696	Hs.35598	ESTs	12.36
	451246	AW189232	Hs.39140	cutaneous T-cell lymphoma tumor antigen	12.36
15	432892	AL042615	Hs.15995	ESTs	12.35
	418982	AI348838	Hs.13073	ESTs	12.35
	414516	AI307802	Hs.279551	ESTs	12.34
	440134	BE410734		gb:601301619F1 NIH_MGC_21 Homo sapiens c	12.29
	443873	AL048542	Hs.16291	ESTs	12.28
	401286				12.26
20	454020	AW962845	Hs.256527	ESTs	12.24
	420077	AW512260	Hs.87767	ESTs	12.24
	443837	AI984625	Hs.9884	spindle pole body protein	12.24
	407519	X64979		gb:H.sapiens mRNA HTPCRX01 for olfactory	12.23
25	435839	AF249744	Hs.25951	Rho guanine nucleotide exchange factor (12.22
	448552	AW973653	Hs.20104	hypothetical protein FLJ00052	12.20
	405325				12.20
	451009	AA013140	Hs.115707	ESTs	12.18
	423066	Y18264	Hs.120171	ESTs	12.17
	439556	AI623752	Hs.163603	ESTs	12.16
30	443062	N77999	Hs.8963	Homo sapiens mRNA full length insert cDN	12.15
	445873	AA250970	Hs.251946	Homo sapiens cDNA: FLJ23107 fis, clone L	12.14
	453542	AW836724	Hs.33190	Homo sapiens mRNA expressed only in plac	12.11
	440106	AA864968	Hs.127699	ESTs	12.10
35	417605	AF006609	Hs.82294	regulator of G-protein signalling 3	12.10
	440286	U29589	Hs.7138	cholinergic receptor, muscarinic 3	12.04
	420061	AW024937	Hs.29410	ESTs	12.02
	458727	AI022813	Hs.92679	Homo sapiens clone CDABP0014 mRNA sequen	11.96
	445407	AI222658	Hs.221889	ESTs, Weakly similar to la costa [D.mela	11.95
40	418250	U29926	Hs.83918	adenosine monophosphate deaminase (isofo	11.94
	414129	AI990287	Hs.270798	ESTs	11.93
	409799	D11928	Hs.76845	phosphoserine phosphatase-like	11.92
	438461	AW075485	Hs.286049	phosphoserine aminotransferase	11.92
	443912	R37257	Hs.184780	ESTs	11.92
45	424606	AA343936		gb:EST49786 Gall bladder I Homo sapiens	11.90
	434217	AW014795	Hs.23349	ESTs	11.90
	451533	NM_004657	Hs.26530	serum deprivation response (phosphatidyl	11.90
	422423	AF283777	Hs.116481	CD72 antigen	11.89
	409398	AW386461		gb:PM4-PT0019-121299-004-F02 PT0019 Homo	11.89
50	423853	AB011537	Hs.133466	slit (Drosophila) homolog 1	11.82
	446180	AI074413	Hs.14220	hypothetical protein FLJ20450	11.80
	414341	D80004	Hs.75909	KIAA0182 protein	11.80
	406538				11.79
	433253	AW450502	Hs.24218	ESTs	11.79
55	447397	BE247676	Hs.18442	E-1 enzyme	11.78
	451684	AF216751	Hs.26813	CDA14	11.76
	416862	R23765	Hs.23575	ESTs	11.74
	425770	NM_014363	Hs.159492	spastic ataxia of Charlevoix-Saguenay (s	11.72
	428826	AL048842	Hs.194019	atractin	11.72
60	433037	NM_014158	Hs.279938	HSPC067 protein	11.72
	447476	BE293466	Hs.20880	ESTs	11.72
	452092	BE245374	Hs.27842	hypothetical protein FLJ11210	11.72
	412922	M60721	Hs.74870	H2.0 (Drosophila)-like homeo box 1	11.72
	401680	NM_005578	Hs.180398	LIM domain-containing preferred transloc	11.69
65	422576	BE548555	Hs.118554	CGI-83 protein	11.68
	450203	AF097994	Hs.301528	L-kynurenine/alpha-aminoadipate aminotra	11.68
	410531	AW752953		gb:QV0-CT0224-261099-035-g02 CT0224 Homo	11.67
	425917	W28517	Hs.117167	Homo sapiens cDNA: FLJ23067 fis, clone L	11.66
	418693	AI750878	Hs.87409	thrombospondin 1	11.64
	400557				11.62

	416188	BE157260	Hs.79070	v-myc avian myelocytomatosis viral oncog	11.60
	419047	AW952771	Hs.90043	ESTs	11.59
	420441	AI986160	Hs.88446	ESTs	11.59
	400885				11.57
5	409853	AW502327		gb:U1-HF-BR0p-aka-a-07-0-U1.r1 NIH_MGC_5	11.56
	400802				11.56
	434540	NM_016045	Hs.5184	TH1 drosophila homolog	11.55
	431449	M55994	Hs.256278	tumor necrosis factor receptor superfami	11.55
10	425928	S55736	Hs.238852	ESTs, Weakly similar to hypothetical pro	11.54
	434701	AA460479	Hs.4096	KIAA0742 protein	11.53
	434228	Z42047	Hs.283978	ESTs; KIAA0738 gene product	11.52
	420729	AW964897	Hs.290825	ESTs	11.52
	428328	AA426080	Hs.98489	ESTs	11.50
	433887	AW204232	Hs.279522	ESTs	11.50
15	414812	X72755	Hs.77367	monokine induced by gamma interferon	11.46
	457718	F18572	Hs.22978	ESTs	11.44
	452260	AA453208	Hs.28726	RAB9, member RAS oncogene family	11.42
	459029	AA131376	Hs.285203	fibroblast growth factor 12	11.42
	456267	AI127958	Hs.83393	cystatin E/M	11.39
20	433285	AW975944	Hs.237396	ESTs	11.38
	449186	AW291876	Hs.196986	ESTs	11.37
	447861	AI434593	Hs.164294	ESTs	11.37
	456023	R00028		gb:ye70a06.s1 Soares fetal liver spleen	11.36
	439444	AI277652	Hs.54578	ESTs	11.31
25	401163				11.31
	430886	L36149	Hs.248116	chemokine (C motif) XC receptor 1	11.28
	450784	AW246803	Hs.47289	ESTs	11.28
	452391	AL044829	Hs.29331	carnitine palmitoyltransferase I, muscle	11.27
	449625	NM_014253	Hs.23796	odc (odd Oz/ten-m, Drosophila) homolog 1	11.26
30	456827	AA075687	Hs.147176	epidermal growth factor receptor substra	11.24
	439328	W07411	Hs.118212	ESTs, Moderately similar to ALU3_HUMAN A	11.24
	432093	H28383		gb:yl52c03.r1 Soares breast 3NbHBst Homo	11.24
	407335	AA631047	Hs.158761	Homo sapiens cDNA FLJ13054 fis, clone NT	11.23
	442501	AA315267	Hs.23128	ESTs	11.22
35	429746	AJ237672	Hs.214142	5,10-methylenetetrahydrofolate reductase	11.21
	422858	R35398		gb:yg64g10.r1 Soares infant brain 1NIB H	11.20
	415156	X84908	Hs.78060	phosphorylase kinase, beta	11.20
	446713	AV660122	Hs.282675	ESTs	11.20
	452221	C21322	Hs.11577	ESTs	11.20
40	418261	W78902	Hs.293297	ESTs	11.17
	433332	AI367347	Hs.127809	ESTs	11.16
	434539	AW748078	Hs.214410	ESTs	11.16
	413471	BE142098		gb:CM4-HT0137-220999-017-d11 HT0137 Homo	11.14
	410037	AB020725	Hs.58009	KIAA0918 protein	11.14
45	405601				11.13
	458332	AI000341	Hs.220491	ESTs	11.12
	427654	AA410183	Hs.137475	ESTs	11.12
	427138	N77624	Hs.173717	phosphatidic acid phosphatase type 2B	11.10
	431475	AI567669	Hs.287316	ESTs	11.10
50	425710	AF030880	Hs.159275	solute carrier family, member 4	11.08
	413748	AW104057	Hs.19193	ESTs	11.07
	409208	Y00093	Hs.51077	integrin, alpha X (antigen CD11C (p150),	11.07
	457278	W92745	Hs.193324	ESTs	11.03
	407021	U52077		gb:Human mariner1 transposase gene, comp	11.02
55	445701	AF055581	Hs.13131	lymphocyte adaptor protein	11.02
	408338	AW867079		gb:MR1-SN0033-120400-002-c10 SN0033 Homo	10.95
	401030	BE382701	Hs.25960	v-myc avian myelocytomatosis viral relat	10.95
	437891	AW006969	Hs.6311	hypothetical protein FLJ20859	10.94
	453874	AW591783	Hs.36131	collagen, type XIV, alpha 1 (undulin)	10.94
60	421562	AA530994	Hs.105803	ghrelin precursor	10.92
	413431	AW246428	Hs.75355	ubiquitin-conjugating enzyme E2N (homolo	10.92
	400132				10.92
	436420	AA443966	Hs.31595	ESTs	10.90
	424880	NM_000328	Hs.153614	retinitis pigmentosa GTPase regulator	10.88
65	433264	D85782	Hs.3229	cysteine dioxygenase, type I	10.88
	429842	AI366213	Hs.173422	KIAA1605 protein	10.87
	412405	AW948126		gb:RC0-MT0013-280300-031-a12 MT0013 Homo	10.85
	400615				10.80
	425018	BE245277	Hs.154196	E4F transcription factor 1	10.80

	456011	BE243628		gb:TCBAP1D1053 Pediatric pre-B cell acut	10.79
	455982	BE176862		gb:RC4-HT0587-170300-012-a04 HT0587 Homo	10.74
	450418	BE218418	Hs.201802	ESTs	10.73
5	412490	AW803564	Hs.288850	ESTs	10.72
	436962	AW377314	Hs.5364	DKFZP564I052 protein	10.70
	437743	AI383497	Hs.131811	ESTs, Weakly similar to ALU1_HUMAN ALU S	10.70
	449967	R40978	Hs.271498	ESTs, Moderately similar to ALU1_HUMAN A	10.70
	449590	AA694070	Hs.268835	ESTs	10.68
10	446035	NM_006558	Hs.13565	Sam68-like phosphotyrosine protein, T-ST	10.68
	426530	U24578	Hs.170250	complement component 4A	10.66
	428600	AW863261	Hs.15036	ESTs, Highly similar to AF161358 1 HSPC0	10.64
	420090	AA220238	Hs.94986	ribonuclease P (38kD)	10.64
	451593	AF151879	Hs.26706	CGI-121 protein	10.62
	438893	AF075031	Hs.29327	ESTs	10.62
15	459324	AW080953		gb:xc28c12.x1 NCI_CGAP_Co18 Homo sapiens	10.61
	439883	AL359652	Hs.171096	Homo sapiens EST from clone DKFZp434A041	10.58
	406513	AA715328	Hs.291205	ESTs	10.57
	407826	AA128423	Hs.40300	calpain 3, (p94)	10.57
	419550	D50918	Hs.90998	KIAA0128 protein; septin 2	10.56
20	428522	R10184	Hs.191987	ESTs, Weakly similar to ALU1_HUMAN ALU S	10.56
	459526	AI142350	Hs.146735	EST	10.55
	411448	AA178955	Hs.271439	ESTs	10.54
	410102	AW248508	Hs.279727	ESTs;	10.52
	406577				10.52
25	408405	AK001332	Hs.44672	hypothetical protein FLJ10470	10.51
	428966	AF059214	Hs.194687	cholesterol 25-hydroxylase	10.50
	400880				10.48
	415875	AA894876	Hs.5687	protein phosphatase 1B (formerly 2C), ma	10.48
30	434715	BE005346	Hs.116410	ESTs	10.46
	406851	AA609784	Hs.180255	major histocompatibility complex, class	10.44
	413409	AI638418	Hs.21745	ESTs	10.44
	418489	U76421	Hs.85302	adenosine deaminase, RNA-specific, B1 (h	10.44
	419465	AW500239	Hs.21187	Homo sapiens cDNA: FLJ23068 fis, clone L	10.44
35	419544	AI909154		gb:QV-BT200-010499-007 BT200 Homo sapien	10.44
	432180	Y18418	Hs.272822	RuvB (E coli homolog)-like 1	10.44
	413822	R08950	Hs.272044	ESTs, Weakly similar to ALU1_HUMAN ALU S	10.42
	437446	AA788946	Hs.16869	ESTs, Moderately similar to CA1C RAT COL	10.41
	415701	NM_003878	Hs.78619	gamma-glutamyl hydrolase (conjugase, fol	10.41
40	443790	NM_003500	Hs.9795	acyl-Coenzyme A oxidase 2, branched chai	10.40
	458873	AW150717	Hs.296176	STAT induced STAT inhibitor 3	10.38
	415082	AA160000	Hs.137396	ESTs	10.37
	429124	AW505086	Hs.196914	minor histocompatibility antigen HA-1	10.36
	417187	AB011151	Hs.81505	KIAA0579 protein	10.34
45	426827	AW067805	Hs.172665	methylenetetrahydrofolate dehydrogenase	10.34
	424280	NM_000030	Hs.271366	alanine-glyoxylate aminotransferase homo	10.33
	446099	T93096	Hs.17126	ESTs	10.32
	423445	NM_014324	Hs.128749	alpha-methylacyl-CoA racemase	10.31
	409995	AW960597	Hs.30164	ESTs	10.30
50	432242	AW022715	Hs.162160	ESTs, Weakly similar to ALU4_HUMAN ALU S	10.30
	406394	AA172106	Hs.110950	Rag C protein	10.30
	406189				10.29
	422283	AW411307	Hs.114311	CDC45 (cell division cycle 45, S.cerevis	10.26
	401598	AA172106	Hs.110950	Rag C protein	10.26
55	456995	T89832	Hs.170278	ESTs	10.26
	416511	NM_006762	Hs.79356	Lysosomal-associated multispinning membr	10.24
	427274	NM_005211	Hs.174142	colony stimulating factor 1 receptor, fo	10.24
	401384				10.23
	456226	D13168	Hs.82002	endothelin receptor type B	10.22
60	426928	AF037062	Hs.172914	retinol dehydrogenase 5 (11-cisand 9-cis	10.21
	423032	AI684746	Hs.119274	ESTs	10.20
	436556	AI364997	Hs.7572	ESTs	10.20
	418400	BE243026	Hs.301989	KIAA0246 protein	10.19
	437401	AA757196	Hs.121190	ESTs	10.19
65	403690				10.17
	423790	BE152393		gb:CM2-HT0323-171199-033-a08 HT0323 Homo	10.16
	434094	AA305599	Hs.238205	hypothetical protein PRO2013	10.16
	434967	AW975009	Hs.292274	ESTs	10.16
	432827	Z68128	Hs.3109	Rho GTPase activating protein 4	10.16
	432660	AI288430	Hs.64004	ESTs	10.14

	452234	AW084176	Hs.223296	ESTs	10.14
	445629	AI245701		gb:qk31f05.x1 NCI_CGAP_Kid3 Homo sapiens	10.13
	457236	AA626142	Hs.179991	ESTs, Weakly similar to KPCE_HUMAN PROTE	10.13
	444605	AI174603	Hs.254105	enolase 1, (alpha)	10.12
5	450313	AI038989	Hs.24809	hypothetical protein FLJ10826	10.12
	407482	NM_006056			10.12
	449971	AA807346	Hs.288581	Homo sapiens cDNA FLJ14296 fis, clone PL	10.11
	441201	AW118822	Hs.128757	ESTs	10.10
	435157	AW014605	Hs.179872	ESTs	10.10
10	417308	H60720	Hs.81892	KIAA0101 gene product	10.09
	442582	AI204266	Hs.179303	ESTs	10.05
	437252	AI433833	Hs.164159	ESTs, Weakly similar to ALU1_HUMAN ALU S	10.04
	448663	BE614599	Hs.106823	H.sapiens gene from PAC 42616, similar t	10.04
	434467	BE552368	Hs.231853	Homo sapiens cDNA FLJ13445 fis, clone PL	10.04
15	423698	AA329796	Hs.1098	DKFZp434J1813 protein	10.02
	412707	AW206373	Hs.16443	Homo sapiens cDNA: FLJ21721 fis, clone C	10.00
	414658	X58528	Hs.76781	ATP-binding cassette, sub-family D (ALD)	10.00
	421832	NM_016098	Hs.108725	HSPC040 protein	10.00
	423554	M90516	Hs.1674	glutamine-fructose-6-phosphate transamin	10.00
20	452039	AI922988	Hs.172510	ESTs	10.00
	434673	AW137442	Hs.136965	ESTs	10.00
	427978	AA418280	Hs.180040	Homo sapiens cDNA: FLJ22439 fis, clone H	10.00
	457803	BE501815	Hs.198011	ESTs	9.99
	428279	AA425310	Hs.155766	ESTs	9.98
25	444412	AI147652	Hs.216381	Homo sapiens clone HH409 unknown mRNA	9.98
	417049	N72394	Hs.44862	ESTs	9.96
	427509	M62505	Hs.2161	complement component 5 receptor 1 (C5a I	9.96
	445424	AB028945	Hs.12696	cortactin SH3 domain-binding protein	9.96
	443678	AW009605	Hs.231923	ESTs	9.96
30	447567	AW474513	Hs.224397	ESTs, Weakly similar to B48013 proline-r	9.94
	414709	AA704703	Hs.77031	Sp2 transcription factor	9.94
	434596	T59538		gb:yb65g12.s1 Stratagene ovary (937217)	9.94
	427630	BE276115	Hs.144980	ESTs, Weakly similar to CA13_HUMAN COLLA	9.93
35	416111	AA033813	Hs.79018	chromatin assembly factor 1, subunit A (9.92
	423349	AF010258	Hs.127428	homeo box A9	9.92
	424308	AW975531	Hs.154443	minichromosome maintenance deficient (S.	9.92
	416814	AW192307	Hs.80042	dolichyl-P-Glc:Man9GlcNAc2-PP-dolichylgl	9.90
	417986	AA481003	Hs.97128	ESTs	9.90
	425174	D87450	Hs.154978	KIAA0261 protein	9.90
40	438171	AW976507	Hs.293515	ESTs	9.90
	421984	AW972187	Hs.110443	hypothetical protein FLJ22215	9.89
	408597	NM_005291	Hs.46453	G protein-coupled receptor 17	9.88
	413907	AI097570	Hs.71222	ESTs	9.87
45	451296	AW801383	Hs.118578	H.sapiens mRNA for ribosomal protein L18	9.86
	433409	AI278802	Hs.25661	ESTs	9.85
	450360	AW117416	Hs.245484	ESTs	9.85
	433104	AL043002	Hs.128246	ESTs, Moderately similar to unnamed prot	9.84
	449824	AI962552	Hs.226765	ESTs	9.84
	452744	AI267652	Hs.30504	Homo sapiens mRNA; cDNA DKFZp434E082 (fr	9.82
50	431066	AF026273	Hs.249175	interleukin-1 receptor-associated kinase	9.82
	426457	AW894667	Hs.169965	chimerin (chimaerin) 1	9.80
	443371	AI792888	Hs.145489	ESTs	9.80
	437159	AL050072		gb:Homo sapiens mRNA; cDNA DKFZp566E1346	9.75
	425242	D13635	Hs.155287	KIAA0010 gene product	9.74
55	447498	N67619	Hs.43687	ESTs	9.74
	426759	AI590401	Hs.21213	ESTs	9.73
	435129	AI381659	Hs.267086	ESTs	9.72
	437672	AW748265	Hs.5741	flavoheprotein b5+b5R	9.72
	438209	AL120659	Hs.6111	KIAA0307 gene product	9.72
60	438440	AA807228	Hs.225161	ESTs	9.72
	449720	AA311152	Hs.288708	ESTs; Weakly similar to KIAA0226 [H.sapi	9.72
	414291	AI289619	Hs.13040	ESTs	9.72
	436206	AK001451	Hs.265561	CD2-associated protein	9.70
	446896	T15767	Hs.22452	Homo sapiens cDNA: FLJ21084 fis, clone C	9.70
65	412667	AW977540	Hs.269254	ESTs	9.70
	423301	S67580	Hs.1645	cytochrome P450, subfamily IVA, polypept	9.67
	440757	AW118645	Hs.160004	ESTs	9.67
	441412	AI393657	Hs.159750	ESTs	9.66
	421044	AF061871	Hs.101302	collagen, type XII, alpha 1	9.66

	414726	BE466863	Hs.280099	ESTs	9.66
	418485	R91679	Hs.124981	ESTs	9.66
	433480	X02422	Hs.181125	immunoglobulin lambda locus	9.65
	441530	AI248301	Hs.127112	ESTs	9.65
5	433533	D53304	Hs.65394	ESTs	9.65
	421470	R27496	Hs.1378	annexin A3	9.64
	438613	C05569	Hs.243122	hypothetical protein FLJ13057 similar to	9.64
	429324	AA488101	Hs.199245	inactivation escape 1	9.62
	450244	AA007534	Hs.125062	ESTs	9.62
10	407660	AW063190	Hs.279101	ESTs	9.61
	406554				9.60
	426404	AA377607	Hs.273138	ESTs	9.58
	447045	AW392394	Hs.278569	KIAA0064 gene product	9.58
	449894	AK001578	Hs.24129	hypothetical protein FLJ10716	9.58
15	448376	AI494332	Hs.196963	ESTs	9.58
	407902	AL117474	Hs.41181	Homo sapiens mRNA; cDNA DKFZp727C191 (fr	9.56
	446572	AV659151	Hs.282961	ESTs	9.56
	459245	BE242623	Hs.31939	manic fringe (Drosophila) homolog	9.55
	423545	AP000692	Hs.129781	chromosome 21 open reading frame 5	9.54
20	414697	BE266134	Hs.76927	translocase of outer mitochondrial membr	9.54
	410846	AW807057		gb:MR4-ST0062-031199-018-b03 ST0062 Homo	9.52
	421181	NM_005574	Hs.184585	LIM domain only 2 (rhombotin-like 1)	9.52
	427308	D26067	Hs.174905	KIAA0033 protein	9.52
25	415995	NM_004573	Hs.994	phospholipase C, beta 2	9.51
	434846	AW295389	Hs.119768	ESTs	9.51
	414342	AA742181	Hs.75912	Homo sapiens cDNA: FLJ22199 fis, clone H	9.50
	416959	D28459	Hs.80612	ubiquitin-conjugating enzyme E2A (RAD6 h	9.50
	443123	AA094538	Hs.6588	ESTs	9.50
30	439312	AA833902	Hs.270745	ESTs	9.48
	449375	R07114	Hs.271224	ESTs	9.48
	436357	AJ132085		gb:Homo sapiens mRNA for axonemal dynein	9.44
	458723	AW137726	Hs.244352	ESTs, Moderately similar to laminin alph	9.44
	457526	AW450584	Hs.192131	ESTs, Weakly similar to RIBB [H.sapiens]	9.43
	404741				9.43
35	422409	NM_005428	Hs.116237	vav 1 oncogene	9.43
	403708				9.42
	408806	AW847814	Hs.289005	Homo sapiens cDNA: FLJ21532 fis, clone C	9.42
	417380	T06809		gb:EST04698 Fetal brain, Stratagene (cat	9.42
40	422501	AA354690	Hs.144967	ESTs	9.42
	426197	AA004410	Hs.167835	acyl-Coenzyme A oxidase 1, palmitoyl	9.42
	452624	AU076606	Hs.30054	coagulation factor V (proaccelerin, labi	9.42
	412110	AW893569		gb:RCO-NN0021-040400-021-c10 NN0021 Homo	9.41
	414158	AA361623	Hs.288775	Homo sapiens cDNA FLJ13900 fis, clone TH	9.41
	408101	AW968504	Hs.123073	CDC2-related protein kinase 7	9.40
45	414171	AA360328	Hs.865	RAP1A, member of RAS oncogene family	9.40
	415947	U04045	Hs.78934	mutS (E. coli) homolog 2 (colon cancer,	9.40
	426959	BE262745		gb:601153869F1 NIH_MGC_19 Homo sapiens c	9.39
	417519	AI689987	Hs.177669	ESTs, Weakly similar to RMS1_HUMAN REGUL	9.39
50	457181	BE514362	Hs.296422	FK506-binding protein 3 (25kD)	9.39
	402835				9.38
	404632				9.38
	446566	H95741	Hs.17914	Homo sapiens cDNA: FLJ22801 fis, clone K	9.37
	455369	AW903533		gb:CM1-NN1031-060400-178-d05 NN1031 Homo	9.37
55	444001	AI095087	Hs.152299	ESTs, Moderately similar to ALU5_HUMAN A	9.36
	458191	AI420611	Hs.127832	ESTs	9.36
	431374	BE258532	Hs.251871	CTP synthase	9.34
	429327	AA283981	Hs.199248	prostaglandin E receptor 4 (subtype EP4)	9.33
	407061	X97748		gb:H.sapiens PTX3 gene promotor region.	9.33
60	416967	BE616731	Hs.80645	interferon regulatory factor 1	9.33
	423013	AW875443	Hs.22209	secreted modular calcium-binding protein	9.33
	439461	AA693960	Hs.103158	ESTs	9.33
	418830	BE513731	Hs.88959	Human DNA sequence from clone 967N21 on	9.32
	422763	AA033699	Hs.83938	ESTs, Moderately similar to MASP-2 [H.sa	9.32
	442739	NM_007274	Hs.8679	cytosolic acyl coenzyme A thioester hydr	9.32
65	452859	AI300555	Hs.288158	Homo sapiens cDNA: FLJ23591 fis, clone L	9.32
	403237				9.32
	415000	AW025529	Hs.239812	ESTs, Weakly similar to CALM_HUMAN CALMO	9.31
	417951	AW976410	Hs.289069	Homo sapiens cDNA: FLJ21016 fis, clone C	9.30
	419066	Z98492	Hs.6975	PRO1073 protein	9.30

	448443	AW167128	Hs.231934	ESTs	9.30
	405125				9.30
	409768	AW499566		gb:UI-HF-BR0p-aji-h-03-0-UI.r1 NIH_MGC_5	9.28
5	453708	AI191811	Hs.54629	ESTs	9.28
	442271	AF000652	Hs.8180	syndecan binding protein (syntenin)	9.27
	410055	AJ250839	Hs.58241	gene for serine/threonine protein kinase	9.26
	448692	AW013907	Hs.224276	ESTs, Moderately similar to predicted us	9.26
	417381	AF164142	Hs.82042	solute carrier family 23 (nucleobase tra	9.25
10	422497	D29642	Hs.1528	KIAA0053 gene product	9.25
	414140	AA281279	Hs.23317	ESTs	9.24
	435980	AF274571	Hs.129142	ESTs; Weakly similar to DEOXYRIBONUCLEAS	9.24
	458530	BE395035	Hs.199889	ESTs, Weakly similar to KIAA0874 protein	9.24
	402585				9.24
15	420819	AA280700		gb:zs95h11.s1 NCL_CGAP_GCB1 Homo sapiens	9.23
	444755	AA431791	Hs.183001	ESTs	9.22
	411630	U42349	Hs.71119	Putative prostate cancer tumor suppresso	9.22
	421246	AW582962	Hs.300961	ESTs, Highly similar to AF151805 1 CGI-4	9.20
	421924	BE514514	Hs.109606	coronin, actin-binding protein, 1A	9.19
20	414888	AL039185	Hs.77558	thyroid hormone receptor interactor 7	9.18
	434267	AI206589	Hs.116243	ESTs	9.17
	409213	U61412	Hs.51133	PTK6 protein tyrosine kinase 6	9.17
	428242	H55709	Hs.2250	leukemia inhibitory factor (cholinergic	9.16
	451736	AW080356	Hs.293684	ESTs, Weakly similar to alternatively sp	9.15
25	413627	BE182082	Hs.246973	ESTs	9.14
	416134	AA528402	Hs.74861	activated RNA polymerase II transcriptio	9.14
	449251	AW151660	Hs.31444	ESTs	9.14
	452813	U54727	Hs.191445	ESTs	9.14
	443622	AI911527	Hs.11805	ESTs	9.14
30	413260	BE075281		gb:PM1-BT0585-290200-005-d07 BT0585 Homo	9.12
	413450	Z99716	Hs.75372	N-acetylgalactosaminidase, alpha-	9.12
	446442	BE221533	Hs.257858	ESTs	9.12
	438540	AA810021	Hs.136906	ESTs	9.12
	426251	M24283	Hs.168383	Intercellular adhesion molecule 1 (CD54)	9.11
35	410290	AA402307	Hs.73818	ubiquinol-cytochrome c reductase hinge p	9.10
	437398	AA913736	Hs.126715	ESTs	9.10
	421559	NM_014720	Hs.105751	Ste20-related serine/threonine kinase	9.10
	439699	AF086534	Hs.187561	ESTs, Moderately similar to ALU1_HUMAN A	9.10
	430799	C19035	Hs.164259	ESTs	9.09
40	424544	M88700	Hs.150403	dopa decarboxylase (aromatic L-amino aci	9.08
	453942	AW190920	Hs.19928	ESTs	9.08
	425844	T68073	Hs.159628	serine (or cysteine) proteinase inhibito	9.08
	434658	AI624436	Hs.194488	ESTs	9.07
	453999	BE328153	Hs.240087	ESTs	9.06
45	436490	R71543	Hs.18713	ESTs	9.05
	409192	AA065131	Hs.233439	ESTs, Weakly similar to ALU7_HUMAN ALU S	9.05
	446223	BE300091	Hs.119699	hypothetical protein FLJ12969	9.04
	447247	AW369351	Hs.287955	Homo sapiens cDNA FLJ13090 fis, clone NT	9.04
	450094	AI174947	Hs.295789	Homo sapiens mRNA; cDNA DKFZp564D1164 (f	9.04
50	432012	AW301344	Hs.195969	ESTs	9.04
	422520	AU076730	Hs.117977	kinesin 2 (60-70kD)	9.02
	418650	BE386750	Hs.86978	prolyl endopeptidase	9.02
	423008	M81590	Hs.123016	5-hydroxytryptamine (serotonin) receptor	9.02
	436476	AA326108	Hs.53631	ESTs	9.02
55	448206	BE622585	Hs.3731	ESTs	9.02
	431574	AW572659	Hs.261373	adenosine A2b receptor pseudogene	9.01
	443453	R99876	Hs.269882	ESTs	9.01
	435472	AW972330	Hs.283022	triggering receptor expressed on myeloid	9.01
	420337	AW295840	Hs.14555	Homo sapiens cDNA: FLJ21513 fis, clone C	9.00
60	449810	AB008681	Hs.23994	activin A receptor, type IIB	9.00
	406780	AA902386	Hs.286	ribosomal protein L4	8.99
	429169	AW341130	Hs.197757	ESTs, Moderately similar to FGFE_HUMAN F	8.99
	421326	AF051428	Hs.103504	estrogen receptor 2 (ER beta)	8.97
	425491	AA883316	Hs.255221	ESTs	8.96
65	425516	BE000707	Hs.29567	ESTs	8.96
	439773	AI051313	Hs.143315	ESTs	8.96
	443247	BE614387	Hs.47378	ESTs	8.96
	456623	AI084125	Hs.108106	transcription factor	8.95
	438707	L08239	Hs.5326	porcupine	8.95
	402240				8.95

	444152	AI125694	Hs.149305	Homo sapiens cDNA FLJ14264 fis, clone PL	8.95
	409842	AW501756		gb:U1-HF-BR0p-ajm-c-09-0-U1.r1 NIH_MGC_5	8.94
	416277	W78765	Hs.73580	ESTs	8.94
5	456697	AI908006	Hs.111334	ferritin, light polypeptide	8.94
	410762	AF226053	Hs.66170	HSKM-B protein	8.92
	412942	AL120344	Hs.75074	mitogen-activated protein kinase-activat	8.92
	442320	AI287817	Hs.129636	ESTs	8.92
	449673	AA002064	Hs.18920	ESTs	8.91
	411486	N85785	Hs.181165	eukaryotic translation elongation factor	8.90
10	437916	BE566249	Hs.20999	Homo sapiens cDNA: FLJ23142 fis, clone L	8.90
	442732	AA257161	Hs.8658	hypothetical protein DKFZp434E0321	8.89
	419741	NM_007019	Hs.93002	ubiquitin carrier protein E2-C	8.89
	411499	AW849292		gb:IL3-CT0215-020300-090-E06 CT0215 Homo	8.89
	431154	AW971228	Hs.290259	ESTs	8.89
15	414922	D00723	Hs.77631	glycine cleavage system protein H (amino	8.88
	418036	Z37976	Hs.83337	latent transforming growth factor beta b	8.87
	406422				8.87
	422926	NM_016102	Hs.121748	ring finger protein 16	8.87
	435220	D50030	Hs.104	HGF activator	8.86
20	418203	X54942	Hs.83758	CDC28 protein kinase 2	8.86
	418613	AA744529	Hs.86575	mitogen-activated protein kinase kinase	8.85
	439250	H66566	Hs.271711	ESTs	8.85
	432359	AA076049	Hs.274415	Homo sapiens cDNA FLJ10229 fis, clone HE	8.84
	450000	AI952797	Hs.10888	Homo sapiens cDNA: FLJ21559 fis, clone C	8.83
25	425657	T89839	Hs.119471	ESTs	8.83
	425694	U51333	Hs.159237	hexokinase 3 (white cell)	8.82
	419972	AL041465	Hs.294038	ESTs, Moderately similar to ALU2_HUMAN A	8.82
	436396	AI683487	Hs.299112	Homo sapiens cDNA FLJ11441 fis, clone HE	8.82
	413413	D82520	Hs.301834	Homo sapiens cDNA FLJ10952 fis, clone PL	8.82
30	428807	AA435997	Hs.104930	ESTs	8.82
	415839	R40611	Hs.137565	ESTs	8.81
	419553	N34145	Hs.250614	ESTs	8.80
	420309	AW043637	Hs.21766	ESTs	8.80
	421863	AI952677	Hs.108972	Homo sapiens mRNA; cDNA DKFZp434P228 (fr	8.80
35	447965	AW292577	Hs.94445	ESTs	8.80
	459172	BE063380		gb:PM0-BT0275-291099-002-g10 BT0275 Homo	8.80
	403259				8.78
	411534	AW850473		gb:IL3-CT0219-280100-061-B11 CT0219 Homo	8.78
	456161	BE264645	Hs.282093	Homo sapiens cDNA: FLJ21918 fis, clone H	8.77
40	413654	AA331881	Hs.75454	peroxiredoxin 3	8.76
	401744				8.76
	425348	AL137477	Hs.155912	cadherin-like 24	8.76
	423396	AI382555	Hs.127950	bromodomain-containing 1	8.75
	450649	NM_001429	Hs.297722	Human DNA sequence from clone RP1-85F18	8.75
45	408331	NM_007240	Hs.44229	dual specificity phosphatase 12	8.74
	423872	AB020316	Hs.134015	uronyl 2-sulfotransferase	8.74
	424906	AI566086	Hs.153716	Homo sapiens mRNA for Hmob33 protein, 3'	8.74
	427596	AA449506	Hs.179765	Homo sapiens mRNA; cDNA DKFZp586H1921 (f	8.73
	432488	AA551010	Hs.216640	ESTs	8.72
50	448980	AL137527	Hs.22703	Homo sapiens mRNA; cDNA DKFZp434P1018 (f	8.72
	429455	AI472111	Hs.292507	ESTs	8.71
	429855	AW385597	Hs.138902	ESTs, Weakly similar to B34087 hypotheti	8.71
	441746	H59955	Hs.127829	ESTs	8.70
	411945	AL033527	Hs.92137	v-myc avian myelocytomatosis viral oncog	8.70
55	413492	D87470	Hs.75400	KIAA0280 protein	8.70
	435706	W31254	Hs.7045	GL004 protein	8.70
	433741	AA609019	Hs.159343	ESTs	8.70
	426340	Z97989	Hs.169370	FYN oncogene related to SRC, FGR, YES	8.69
	422779	AA317036	Hs.41989	ESTs	8.67
60	449785	AI225235	Hs.288300	Homo sapiens cDNA: FLJ23231 fis, clone C	8.67
	420144	AA811813	Hs.119421	ESTs	8.66
	420235	AA256756	Hs.31178	ESTs	8.66
	432606	NM_002104	Hs.3066	granzyme K (serine protease, granzyme 3;	8.66
	425762	BE244076	Hs.159578	Homo sapiens mRNA for FLJ00020 protein,	8.65
65	427448	BE246449	Hs.2157	Wiskott-Aldrich syndrome (eczema-thrombo	8.64
	418033	W68180	Hs.259855	Homo sapiens cDNA FLJ12507 fis, clone NT	8.64
	429084	AJ001443	Hs.195614	splicing factor 3b, subunit 3, 130kD	8.64
	417094	NM_006895	Hs.81182	histamine N-methyltransferase	8.64
	457277	NM_004736	Hs.227656	xenotropic and polytropic retrovirus rec	8.63

	422631	BE218919	Hs.118793	hypothetical protein FLJ10688	8.63
	410679	AW795196	Hs.215857	ring finger protein 14	8.63
	431585	BE242803	Hs.262823	hypothetical protein FLJ10326	8.62
	401851				8.62
5	401866				8.62
	407783	AW996872	Hs.172028	a disintegrin and metalloproteinase doma	8.62
	408242	AA251594	Hs.43913	PIBF1 gene product	8.62
	422250	AW408530	Hs.113823	ClpX (caseinolytic protease X, E. coli)	8.62
	430259	BE550182	Hs.127826	RalGEF-like protein 3, mouse homolog	8.62
10	452598	AI831594	Hs.68647	ESTs, Weakly similar to ALU7_HUMAN ALU S	8.62
	419541	AW749617		gb:RC3-BT0502-130100-012-g07 BT0502 Homo	8.60
	428839	AI767756	Hs.82302	ESTs	8.60
	429328	AA829402	Hs.47939	ESTs	8.60
	451491	AI972094	Hs.286221	Homo sapiens cDNA FLJ13741 fis, clone PL	8.60
15	452561	AI692181	Hs.49169	KIAA1634 protein	8.60
	420027	AF009746	Hs.94395	ATP-binding cassette, sub-family D (ALD)	8.60
	435205	X54136	Hs.181125	immunoglobulin lambda locus	8.60
	430900	U91939	Hs.248123	G protein-coupled receptor 25	8.60
	405074				8.59
20	437991	AI479773	Hs.181679	ESTs	8.59
	436346	BE328882	Hs.193096	ESTs, Moderately similar to U119_HUMAN U	8.58
	411079	AA091228		gb:cchn2152.seq.F Human fetal heart, Lam	8.57
	418452	BE379749	Hs.85201	C-type (calcium dependent, carbohydrate-	8.56
	429109	AL008637	Hs.196352	neutrophil cytosolic factor 4 (40kD)	8.56
25	448019	AW947164	Hs.195641	ESTs	8.56
	449865	AW204272	Hs.199371	ESTs	8.55
	431180	H55883		gb:yq94h03.r1 Soares fetal liver spleen	8.54
	445988	BE007663	Hs.13503	inactivation escape 2	8.54
	405876				8.54
30	407235	D20569	Hs.169407	SAC2 (suppressor of actin mutations 2, y	8.54
	414807	AI738616	Hs.77348	hydroxyprostaglandin dehydrogenase 15-(N	8.54
	425671	AF193612	Hs.159142	lunatic fringe (Drosophila) homolog	8.54
	452413	AW082633	Hs.212715	ESTs	8.54
	421620	AA446183	Hs.91885	ESTs	8.53
35	444539	AI955765	Hs.146907	ESTs	8.52
	415102	M31899	Hs.77929	excision repair cross-complementing rode	8.51
	405552				8.51
	418068	AW971155	Hs.293902	ESTs, Weakly similar to prolyl 4-hydroxy	8.50
	420133	AA426117	Hs.14373	ESTs	8.50
40	438887	R68857	Hs.265499	ESTs	8.50
	446468	AI765890	Hs.16341	ESTs; Moderately similar to !!!! ALU SUB	8.50
	446585	AV659397	Hs.282948	ESTs	8.50
	441896	AW891873		gb:CM3-NT0090-040500-173-b02 NT0090 Homo	8.50
	437718	AI927288	Hs.196779	ESTs	8.48
45	420656	AA279098	Hs.187636	ESTs	8.48
	429303	AW137635	Hs.44238	ESTs	8.48
	450624	AL043983	Hs.125063	Homo sapiens cDNA FLJ13825 fis, clone TH	8.48
	452573	AI907957	Hs.287622	Homo sapiens cDNA FLJ14082 fis, clone HE	8.48
	456341	AA229126	Hs.122647	N-myristoyltransferase 2	8.48
50	423024	AA593731	Hs.75613	CD36 antigen (collagen type I receptor,	8.47
	446985	AL038704	Hs.156827	ESTs, Weakly similar to ALU1_HUMAN ALU S	8.46
	431778	AL080276	Hs.268562	regulator of G-protein signalling 17	8.46
	400268				8.46
	421828	AW891965	Hs.289109	dimethylarginine dimethylaminohydrolase	8.45
55	417022	NM_014737	Hs.80905	Ras association (RalGDS/AF-6) domain fam	8.44
	421029	AW057782	Hs.293053	ESTs	8.44
	425171	AW732240	Hs.300615	ESTs	8.44
	459070	AI814302		gb:wj71c12.x1 NCI_CGAP_Lu19 Homo sapiens	8.42
	406006				8.42
60	412643	AW971239	Hs.293982	ESTs	8.42
	424775	AB014540	Hs.153026	SWAP-70 protein	8.42
	446848	AW136083	Hs.195266	ESTs, Weakly similar to S59501 interfero	8.42
	448043	AI458653	Hs.201881	ESTs	8.41
	407183	AA358015		gb:EST66864 Fetal lung III Homo sapiens	8.40
65	412324	AW978439	Hs.69504	ESTs	8.40
	419594	AA013051	Hs.91417	topoisomerase (DNA) II binding protein	8.40
	430968	AW972830		gb:EST384925 MAGE resequences, MAGL Homo	8.40
	431689	AA305688	Hs.267695	UDP-Gal:betaGlcNAc beta 1,3-galactosyltr	8.40
	438582	AI521310	Hs.283365	ESTs, Weakly similar to ALU5_HUMAN ALU S	8.40

	447685	AL122043	Hs.19221	hypothetical protein DKFZp566G1424	8.40
	459119	AW844498	Hs.289052	Homo sapiens LENG8 mRNA, variant C, part	8.38
	400817				8.37
5	425265	BE245297		gb:TCBAP1E2482 Pediatric pre-B cell acut	8.37
	409385	AA071267		gb:zm61g01.r1 Stratagene fibroblast (937	8.36
	439121	BE047779	Hs.44701	ESTs	8.36
	419968	X04430	Hs.93913	interleukin 6 (interferon, beta 2)	8.36
	408327	AW182309	Hs.249963	ESTs, Highly similar to dJ1170K4.4 [H.sa	8.35
	403976				8.34
10	448064	AA379036		gb:EST91809 Synovial sarcoma Homo sapien	8.33
	442914	AW188551	Hs.99519	Homo sapiens cDNA FLJ14007 fis, clone Y7	8.33
	428032	AW997704	Hs.11493	Homo sapiens cDNA FLJ13536 fis, clone PL	8.32
	434194	AF119847	Hs.283940	Homo sapiens PRO1550 mRNA, partial cds	8.32
	458677	AW937670	Hs.254379	ESTs	8.32
15	420925	NM_015698	Hs.100391	T54 protein	8.30
	416475	T70298		gb:yd26g02.s1 Soares fetal liver spleen	8.30
	416852	AF283776	Hs.80285	Homo sapiens mRNA; cDNA DKFZp586C1723 (f	8.30
	430676	AF084866		gb:Homo sapiens envelope protein RIC-3 (8.30
	428455	AI732694	Hs.98520	ESTs	8.29
20	435343	AW194962	Hs.199028	ESTs	8.29
	450783	BE266695		gb:601190242F1 NIH_MGC_7 Homo sapiens cD	8.29
	404946				8.28
	422942	AF054839	Hs.122540	tetraspan 2	8.28
25	453716	AA037675	Hs.152675	ESTs	8.28
	437098	AA744488	Hs.132842	ESTs, Moderately similar to ALU1_HUMAN A	8.28
	443907	AJ076484	Hs.9963	TYRO protein tyrosine kinase binding pro	8.27
	401930	AF106069	Hs.23168	ubiquitin specific protease 15	8.26
	446554	AA151730	Hs.301789	ESTs, Weakly similar to similar to C.ele	8.26
	426290	AB007918	Hs.169182	KIAA0449 protein	8.25
30	419904	AA974411	Hs.18672	ESTs	8.25
	413886	AW958264	Hs.103832	ESTs, Weakly similar to TRHY_HUMAN TRICH	8.24
	424738	AI963740	Hs.46826	ESTs	8.24
	427359	AW020782	Hs.79881	Homo sapiens cDNA: FLJ23006 fis, clone L	8.24
	424534	D87682	Hs.150275	KIAA0241 protein	8.24
35	424429	U63830	Hs.146847	TRAF family member-associated NFKB activ	8.24
	442604	BE263710	Hs.279904	ESTs	8.22
	442992	AI914699	Hs.13297	ESTs	8.22
	427210	BE396283	Hs.173987	eukaryotic translation initiation factor	8.22
	457229	BE222450	Hs.266390	ESTs	8.21
40	423730	AA330214		gb:EST33935 Embryo, 12 week II Homo sapi	8.21
	411928	AA888624	Hs.19121	adaptor-related protein complex 2, alpha	8.20
	416051	AA835868	Hs.25253	Homo sapiens cDNA: FLJ20935 fis, clone A	8.20
	417231	R40739	Hs.21326	ESTs	8.20
45	422049	W25760	Hs.77631	glycine cleavage system protein H (amino	8.20
	427528	AJ077143	Hs.179565	minichromosome maintenance deficient (S.	8.20
	458776	AV654978	Hs.19904	cystathionase (cystathionine gamma-lyase	8.19
	417687	AI828596	Hs.250691	ESTs	8.18
	423218	NM_015896	Hs.167380	BLu protein	8.18
50	425397	J04088	Hs.156346	topoisomerase (DNA) II alpha (170kD)	8.18
	406964	M21305	Hs.247946	Human alpha satellite and satellite 3 ju	8.18
	402401	U42349	Hs.71119	Putative prostate cancer tumor suppresso	8.18
	423397	NM_001838	Hs.1652	chemokine (C-C motif) receptor 7	8.18
	427857	AL133017	Hs.2210	thyroid hormone receptor interactor 3	8.17
	401519				8.17
55	447188	H65423	Hs.17631	Homo sapiens cDNA FLJ20118 fis, clone CO	8.16
	424704	AI263293	Hs.152096	cytochrome P450, subfamily IIJ (arachido	8.16
	435854	AJ278120	Hs.4996	DKFZP564D166 protein	8.14
	448556	AW885606	Hs.5064	ESTs	8.14
	449217	AA278536	Hs.23262	ribonuclease, RNase A family, k6	8.14
60	453124	AI139058	Hs.23296	ESTs	8.14
	442812	AI018406	Hs.131284	ESTs	8.14
	421129	BE439899	Hs.89271	ESTs	8.14

TABLE 9A shows the accession numbers for those primekeys lacking a unigeneID in Table 9. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

10	Pkey:	Unique Eos	probeset identifier number
	CAT number:	Gene cluster number	
	Accession:	Genbank accession numbers	
15	Pkey	CAT number	Accession
	408057	1035720_-1	AW139565
	408069	103655_1	H81795 Z42291 R20973 AA046920
	408182	104479_1	AA047854 AA057506 AA053841
20	408338	1052148_1	AW867079 AW867086 AW182772
	408828	108463_1	BE540279 AW410659 AA057857 R77693 BE278674
	409126	110159_1	AA063426 AW962323 AW408063 AA063503 AA772927 AW753492 BE175371 AA311147
	409292	111586_1	AA071051 AA070584 AA069938 AA102136 AA074430
	409314	111841_1	AA070266 AA084967 AA126998
25	409385	112523_1	AA071267 T65940 T64515 AA071334
	409398	1126716_1	AW386461 AW876408 AW386672 AW386599 AW876258 AW386619 AW386289 AW876136 AW876203 AW876213 AW876301
			AW876295 AW876349 AW876365 AW876160 AW876369 AW876352 AW876271
	409671	114731_1	AA076769 AA076781 AI087968
	409768	1154035_1	AW499566 AW502378 AW499522 AW502046 AW502671 AW501917 AW501868 AW501721 AW502813
30	409841	1156088_1	AW502139 AW502432 AW502235 AW501683 AW502647
	409842	1156119_1	AW501756 AW502096 AW502465 AW501715
	409853	1156226_1	AW502327 AW502488 AW501829 AW502625 AW502687
	410531	1207200_1	AW752953 H88044 BE156092
	410688	1216101_1	AW796342 AW796356 BE161430
35	410846	1223902_1	AW807057 AW807054 AW807189 AW807193 AW807369 AW807429 AW807364 AW807365 AW807078 AW807256 AW807180
			AW807331
	410896	1226053_1	AW809637 AW809697 AW810554 AW809707 AW809885 AW810000 AW810088 AW809742 AW809816 AW809749 AW809639
			AW809722 AW809836 AW809774 AW810023 AW810013 AW809813 AW809660 AW809728 AW809768 AW809951 AW809657
			AW809954
40	411079	123128_1	AA091228 H71860 H71073
	411424	1245497_1	AW845985 AW845991 AW845962
	411499	1248105_1	AW849292 AW849431 AW849422 AW849428 AW849420 AW849424 AW849427
	411507	1248607_1	AW850140 AW850195 AW850192
	411534	1248827_1	AW850473 AW850471 AW850431 AW850523
45	411972	1268491_1	BE074959 AW880160
	412110	1277844_1	AW893569 AW893571 AW893588 AW893593
	412226	1284289_1	W26786 AW998612 AW902272
	412257	1285376_1	AW903830 BE071916
50	412405	1293012_1	AW948126 AW948139 AW948196 AW948145 AW948162 AW948134 AW948127 AW948124 AW948153 AW948157 AW948125
			AW948131 AW948158 AW948164 AW948151
	413260	1356003_1	BE075281 BE075219 BE075123 BE075119 BE075046
	413471	1371778_1	BE142098 BE142092
	413729	1385114_1	BE159999 BE160056 BE160107 BE160139
	414182	142409_1	AA136301 AI381776 AA136321
55	414989	1511339_1	T81668 C19040 C17569
	415354	1534763_1	F06495 R24336 R13046
	416011	1566439_1	H14487 R50911 Z43216
	416475	1596398_1	T70298 H58072 R02750
	417380	1672461_1	T06809 N75735
60	419392	1843934_-1	W28573
	419541	185724_1	AW749617 R64714 AA244138 AA244137 BE094019
	419544	185760_2	AI909154 AA526337 AA244193 AI909153
	420819	196721_1	AA280700 AW975494 AA687385
	421245	200620_1	AA285363 AA285333 AA285359 AA285326 AA285350
65	422673	219674_1	N59027 AA314694 N53937 R08100

	422695	219996_1	AA315158 AW961298 N76067 AW802759 AI858495 W04474
	422858	222209_1	R35398 BE252178 AA318153
	422940	223106_1	BE077458 AA337277 AA319285
5	423730	231462_1	AA330214 AW962519 T54709
	423790	232031_1	BE152393 AA330984 BE073904
	424385	238731_1	AA339666 AW952809 AA349119
	424606	241409_1	AA343936 AA344060 AW963081
	425265	249175_1	BE245297 AA353976 AW505023
	426959	273830_-1	BE262745
10	430676	32168_1	AF084866 AF084870 AF084864 AF084867 AF084869 AF084868 AW818206 AW812038 BE144813 BE144812 AW812041 AW812040 AW812067 BE061583 BE061604 T05808 AI352469 AA580921 BE141783 BE141782 BE061601 AW814393 AW885029
	430968	326269_1	AW972830 AA527647 AA489820 AA570362
	431180	328906_1	H55883 AW971249 AA493900 H55788
15	432093	341283_1	H28383 AW972670 H28359 AA525808
	434596	38937_1	T59538 T59589 T59598 T59542 AF147374
	436357	41842_1	AJ132085 Z83805
	437159	43393_1	AL050072 AW900148
	437495	43765_1	BE177778 BE177779 AL390180 AA359908
20	439097	46858_1	H66948 AF085954 H66949
	439120	46879_1	H56389 AF085977 H56173
	440134	48675_1	BE410734 BE560117 BE270054 BE296330 BE267957 AI003007 BE545259
	441896	52842_1	AW891873 AW891897 BE564764
	445629	645767_1	AI245701 BE272724
25	447229	71288_1	BE617135 AW504051 AW504283
	448064	74761_1	AA379036 AA150589 AI696854 BE621316
	450783	84655_1	BE266695 BE265474 N53200 BE267333
	451045	85673_1	AA215672 AI696628 AA013335 H86334 AA017006
	452549	921802_1	AI907039 AI907081
30	452560	922216_1	BE077084 AW139963 AW863127 AW806209 AW806204 AW806205 AW806206 AW806211 AW806212 AW806207 AW806208 AW806210 AI907497
	452712	928309_1	AW838616 AW838660 BE144343 AI914520 AW888910 BE184854 BE184784
	453758	980026_1	U83527 AL120938 U83522
35	454093	1007366_1	AW860158 AW862385 AW860159 AW862386 AW862341 AW821869 AW821893 AW062660 AW062656
	454563	1224342_1	AW807530 AW807540 AW807537 AW846086 BE141634 AW846089 AW807499 AW807533 AW838499
	454791	1234759_1	BE071874 BE071882 AW820782 AW821007
	454977	1247099_1	AW848032 AW848630 AW848478 AW848623 AW848484 AW848169 AW848830 AW848149 AW848119 AW848893 AW848903 AW848407
40	455131	1254674_1	AW857913 AW857916 AW857914 AW861627 AW861626 AW861624
	455183	1259023_1	AW984111 AW863918 AW863856
	455254	1266449_1	AW877015 AW877133 AW876978 AW877071 AW876988 AW877069 AW877063 AW877013
	455369	1285173_16	AW903533 AW903516 AW903562 BE085202 BE085215 BE085214 BE085209 BE085172 BE085175 BE085193 BE085211 BE085199
45	455982	1396849_1	BE176862 BE176876 BE176947 BE176878
	456011	1410860_1	BE243628 BE246081 BE247016 BE241984 BE241534 BE246091 BE245679 BE243620 BE245998 BE242329 BE241417 BE241457 BE242522 BE241989 BE241464
	456023	1416335_1	R00028 BE247630
	457586	360505_1	AW062439 AW751554 AA579463
50	457595	364225_-1	AA584854
	457751	399422_1	AI908236 AA663731
	459070	883688_1	AI814302 AI814428
	459081	889426_1	W07808 AI822066
	459145	918957_1	AI903354 AI903489 AI903488
	459172	921149_1	BE063380 BE063346 AI906097
55	459234	945240_-1	AI940425

TABLE 9B shows the genomic positioning for those primekeys lacking unigene ID's and accession numbers in Table 9. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique number corresponding to an Eos probeset
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
 Strand: Indicates DNA strand from which exons were predicted.
 Nt_position: Indicates nucleotide positions of predicted exons.

	Pkey	Ref	Strand	Nt_position
15	400452	8113550	Minus	90308-90505
	400557	9801261	Plus	208453-208528,209633-209813
	400615	9908994	Plus	118036-118166,118681-118807
	400802	8567867	Minus	174571-174856
20	400817	8569994	Plus	170793-170948
	400880	9931121	Plus	29235-29336,36363-36580
	400885	9958187	Minus	58242-58733
	400926	7651921	Minus	52033-52158,53956-54120,54957-55052,55420-55480,56452-56666,57221-57718
	400952	7658481	Plus	192667-192826,194387-194876
25	400991	8096825	Plus	159197-159320
	401044	8117619	Plus	73501-73674
	401124	8570296	Minus	124181-124391
	401163	6981820	Plus	5302-5545
	401201	9743387	Minus	138534-138629,139234-139294,140121-140335,142033-142479
30	401286	9801342	Minus	147036-147318
	401384	6850939	Minus	58360-58545
	401468	6433826	Plus	13056-13482
	401515	7630851	Plus	29929-30126
	401519	6649315	Plus	157315-157950
	401672	9838136	Plus	128526-128704,130755-130860
35	401744	2576349	Plus	14595-14751
	401851	7770425	Minus	146443-146664,147794-147971,148351-148480,148980-149111,149801-149949
	401866	8018106	Plus	73126-73623
	402240	7690131	Plus	104382-104527,106136-106372
40	402359	9211204	Minus	40403-41961
	402585	9908890	Minus	174893-175050,183210-183435
	402788	9796102	Plus	98273-101430
	402802	3287156	Minus	53242-53432
	402812	6010110	Plus	25026-25091,25844-25920
	402828	8918414	Plus	69071-69642
45	402835	9187337	Plus	26961-27101
	402838	9369121	Minus	32589-32735,35478-35666
	402842	9369121	Minus	76355-76479
	402895	9967547	Plus	85537-85671,86379-86469
50	402964	9581599	Minus	46624-46784
	403137	9211494	Minus	92349-92572,92958-93084,93579-93712,93949-94072,94591-94748,95214-95337
	403237	7637807	Plus	7271-7527
	403259	7770585	Plus	4693-4857
	403683	7331517	Plus	217175-217446
55	403690	7387384	Minus	78627-79583
	403708	5705981	Minus	134394-134812
	403838	4176355	Plus	19197-19502
	403851	7708872	Plus	22733-23007
	403976	7657840	Plus	24755-24969
60	404407	7329316	Minus	48154-48499
	404426	7407959	Plus	77842-77954
	404632	9796668	Plus	45096-45229
	404741	8574139	Plus	143025-143467
	404756	7706327	Plus	82849-83627
65	404946	7382189	Plus	134445-134750
	405074	7770440	Plus	44340-44559,44790-45059
	405125	8247873	Plus	137113-137814
	405172	9966752	Plus	153027-153262

	405236	7249076	Minus	151699-151915
	405325	6094661	Minus	25818-26380
	405411	3451356	Minus	17503-17778,18021-18290
5	405495	8050952	Minus	72182-72373
	405552	1552506	Plus	45199-45647
	405601	5815493	Minus	147835-147935,149220-149299
	405685	4508129	Minus	37956-38097
	405777	7263187	Minus	104773-105051
10	405856	7653009	Plus	101777-102043
	405876	6758747	Plus	39694-40031
	405932	7767812	Minus	123525-123713
	405934	6758795	Plus	159913-160605
	406006	8247801	Minus	42640-42776
15	406134	9163473	Plus	153291-153452
	406189	7289992	Minus	22007-22234
	406422	9256411	Plus	163003-163311
	406516	7711422	Minus	128375-128449,128560-128784
	406538	7711478	Plus	35196-35367,38229-38476,40080-40216,43522-43840
	406554	7711566	Plus	106956-107121
20	406577	7711730	Plus	11377-11509

TABLE 10: shows genes, including expression sequence tags differentially expressed in taxol resistant prostate tumor xenografts as compared to taxol sensitive prostate tumor xenografts. The genes are indicated as either being upregulated or downregulated during the induction of taxol resistance in sequential passages of the grafts.

10	Pkey:	Unique Eos probeset identifier number															
	ExAccn:	Exemplar Accession number, Genbank accession number															
	UnigeneID:	Unigene number															
	Unigene Title:	Unigene gene title															
	Eos:	Internal Eos name															
15	F00-F14:	passage number															
	Pkey	ExAccn	UnigeneID	UnigenTitle	Eos	Resp.F00	F00	F02	F02	F05	F05	F07	F09	F10	F11	F13	F14
20	117921	N51002	Hs.47170	Liprin A2	PM28UP	1	9	8	9	32	20	34	122	105	82	71	111
	112971	T17185	Hs.4299	ESTs	CHA1down	290	281	267	335	270	284	150	157	83	89	49	75
	126645	AI167942	Hs.61635	STEAP	PAA5down	106	111	103	71	34	67	33	14	2	1	1	1
	119018	N95796	Hs.179809	ESTs	PAB2down	765	841	757	909	742	704	478	428	253	175	228	238
	110844	N31952	Hs.167531	ESTs	PAV7down	175	192	147	141	123	129	73	65	55	48	54	84
25	100654	HG2841-HT2969	Hs.75442	Albumin, A	PM01down	666	605	504	728	357	445	602	187	117	127	117	113
	100655	HG2841-HT2970	Hs.75442	Albumin, A	PM02down	620	653	486	688	368	386	606	175	101	95	115	97
	102076	U09579	Hs.252437	cyclin-dep	PM03down	101	94	143	190	105	107	88	40	34	31	46	22
	102208	U22961	Hs.75442	albumin	PM04down	495	424	323	518	252	296	467	188	169	143	165	145
	103739	AA075779	--	mitochondr	PM05down	75	190	606	230	378	106	218	88	69	192	69	99
30	107036	AA599690	Hs.15725	SBB148	PM06down	87	124	115	188	132	111	66	71	49	70	38	50
	108242	AA062746	--	ESTs	PM07down	14	20	252	13	22	43	193	10	10	104	21	18
	108282	AA065143	--	solute car	PM08down	27	54	178	73	108	37	53	24	14	53	15	34
	108679	AA115963	--	beta-1-glo	PM09down	680	893	1292	656	869	389	1	74	118	662	359	409
	108731	AA126313	Hs.107476	ATP syntha	PM10down	10	19	185	25	60	1	32	3	7	14	1	1
35	110675	H89355	Hs.6598	adrenergic	PM11down	207	334	237	239	231	220	119	145	93	64	56	124
	115412	AA283804	Hs.193552	ESTs	PM12down	146	316	282	271	340	334	115	238	100	196	83	207
	115844	AA430124	Hs.234607	MDM2	PM13down	49	93	94	154	132	91	23	54	23	76	14	41
	120588	AA281591	Hs.16193	ESTs	PM14down	80	157	58	141	159	127	39	83	35	37	16	46
	132349	Y00705	Hs.181286	serine pro	PM15down	146	217	214	150	106	128	177	85	54	63	66	56
40	132888	AA490775	Hs.5920	N-acetylima	PM16down	92	150	132	178	126	139	53	94	48	67	41	80
	132967	AA032221	Hs.61635	STEAP	PM17down	224	208	203	215	205	180	132	65	68	50	48	63
	133063	AA283085	Hs.64065	ESTs	PM18down	85	148	161	150	92	108	42	99	42	65	29	126
	134374	D62633	Hs.8236	ESTs	PM19down	230	240	194	212	231	189	89	123	107	95	68	91
45	135400	M23263	Hs.99915	androgen r	PM20down	36	167	99	178	132	101	23	71	26	122	14	44

TABLE 11: shows genes, including expression sequence tags that are up-regulated in prostate tumor tissue compared to normal prostate tissue as analyzed using Affymetrix/Eos Hu01 GeneChip array. Shown are the ratios of "average" normal prostate to "average" prostate cancer tissues.

5

Pkey: Unique Eos probeset identifier number
 ExAccn: Exemplar Accession number, Genbank accession number
 UnigenelD: Unigene number
 Unigene Title: Unigene gene title
 R1: Background subtracted normal prostate : prostate tumor tissue

	Pkey	ExAccn	UnigenelD	Unigene Title	R1
10	101336	L49169	Hs.75678	FBJ murine osteosarcoma viral oncogene homolog B	0.012
	130642	M63438	Hs.156110	Immunoglobulin kappa variable 1D-8	0.015
	133512	X01677	Hs.195188	glyceraldehyde-3-phosphate dehydrogenase	0.017
	133436	H44631	Hs.737	immediate early protein	0.017
	129292	X13810	Hs.1101	POU domain; class 2; transcription factor 2	0.019
15	100610	HG2566-HT4792		Microtubule-Associated Protein Tau, Alt. Splice 3, Exon 8	0.02
	133448	M34516	Hs.170116	immunoglobulin lambda-like polypeptide 3	0.021
	125193	W67577	Hs.84298	CD74 antigen (invariant polypeptide of major histocompatibility complex; class II antigen-associated)	0.022
20	133456	T49257	Hs.183704	ubiquitin C	0.022
	134546	AA459310	Hs.8518	Homo sapiens mRNA; cDNA DKFZp586L1722 (from clone DKFZp586L1722)	0.023
	102131	U15085	Hs.1162	major histocompatibility complex; class II; DM beta	0.023
	101375	M13560	Hs.84298	CD74 antigen (invariant polypeptide of major histocompatibility complex; class II antigen-associated)	0.023
	100674	HG3033-HT3194		Spliceosomal Protein Sap 62	0.024
25	134365	R32377	Hs.82240	syntaxin 3A	0.027
	132335	D60387	Hs.189885	ESTs	0.027
	110303	H37901	Hs.32706	ESTs	0.028
	131678	N59162	Hs.30542	ESTs	0.028
	116599	D80046	Hs.250879	ESTs	0.029
30	133769	M17733	Hs.75968	thymosin; beta 4; X chromosome	0.029
	107904	AA026648	Hs.61389	ESTs	0.03
	129427	T80746	Hs.111334	ferritin; light polypeptide	0.03
	105987	AA406631	Hs.110299	mitogen-activated protein kinase kinase 7	0.03
	131466	F03233	Hs.27189	ESTs	0.032
35	102859	X00274	Hs.76807	Human HLA-DR alpha-chain mRNA	0.032
	134626	S82198	Hs.8709	caldecrin (serum calcium decreasing factor; elastase IV)	0.032
	134170	M63138	Hs.79572	cathepsin D (lysosomal aspartyl protease)	0.033
	131713	X57809	Hs.181125	immunoglobulin lambda gene cluster	0.034
	100748	HG3517-HT3711		Alpha-1-Antitrypsin, 5' End	0.034
40	118769	N74496		ESTs	0.034
	111734	R25375	Hs.126916	ESTs	0.036
	109221	AA192755	Hs.85840	ESTs; Weakly similar to stac [H.sapiens]	0.036
	133846	AA480073	Hs.76719	U6 snRNA-associated Sm-like protein	0.036
	135281	AA401575	Hs.97757	ESTs	0.037
45	119073	R32894	Hs.45514	v-ets avian erythroblastosis virus E26 oncogene related	0.037
	100760	HG3576-HT3779		Major Histocompatibility Complex, Class II Beta W52	0.037
	101426	M19483	Hs.25	ATP synthase; H+ transprng; mitochondrl F1 complex; beta polypept	0.038
	129568	AA428025	Hs.114360	transforming growth factor beta-stimulated protein TSC-22	0.038
	130900	Z38468	Hs.21036	ESTs; Moderately similar to F25965_3 [H.sapiens]	0.039
50	133879	M13829	Hs.77183	v-raf murine sarcoma 3611 viral oncogene homolog 1	0.039
	100627	HG2702-HT2798		Serine/Threonine Kinase (Gb:Z25424)	0.039
	129424	M55593	Hs.111301	matrix metalloproteinase 2 (gelatinase A; 72kD gelatinase; 72kD type IV collagenase)	0.039
	128652	AA621245	Hs.103147	ESTs; Weakly similar to similar to SP:YR40_BACSU [C.elegans]	0.039
	129979	T72635	Hs.13956	ESTs	0.039
55	133468	X03068	Hs.73931	major histocompatibility complex; class II; DQ beta 1	0.04
	102636	U67092		Human ataxia-telangiectasia locus protein (ATM) gene, exons 1a, 1b, 2, 3 and 4, partial cds	0.04
	129536	M33493	Hs.184504	tryptase; alpha	0.04
	133599	M64788	Hs.75151	RAP1; GTPase activating protein 1	0.041

5	102104	U12139		Human alpha1(XI) collagen (COL11A1) gene, 5' region and exon 1	0.041
	131340	AA478305	Hs.25817	Homo sapiens chromosome 19; cosmid R27216	0.041
	130446	X79510	Hs.155693	protein tyrosine phosphatase; non-receptor type 21	0.042
	101352	L77701	Hs.16297	COX17 (yeast) homolog; cytochrome c oxidase assembly protein	0.042
	122593	AA453310	Hs.128749	alpha-methylacyl-CoA racemase	0.042
10	130181	R39552	Hs.151608	Homo sapiens clone 23622 mRNA sequence	0.042
	134071	Z14093	Hs.78950	branched chain keto acid dehydrogenase E1; alpha polypeptide (maple syrup urine disease)	0.042
	108129	AA053252	Hs.185848	ESTs; Weakly similar to !! ALU SUBFAMILY J WARNING ENTRY !! [H.sapiens]	0.043
	130511	L32137	Hs.1584	cartilage oligomeric matrix protein (pseudoachondroplasia; epiphyseal dysplasia 1; multiple)	0.043
	133336	AA291456	Hs.71190	ESTs	0.043
15	132982	L02326	Hs.198118	immunoglobulin lambda-like polypeptide 2	0.044
	131880	AA047034	Hs.33818	RecQ protein-like 5	0.044
	130540	U35234	Hs.159534	protein tyrosine phosphatase; receptor type; S	0.044
	133467	AA258595	Hs.73931	major histocompatibility complex; class II; DQ beta 1	0.044
	101191	L20688	Hs.83656	Rho GDP dissociation inhibitor (GDI) beta	0.044
20	101860	M95610	Hs.37165	collagen; type IX; alpha 2	0.044
	102799	U88898		Human endogenous retroviral H protease/integrase-derived ORF1 mRNA, complete cds, and putative envelope prot mRNA, partial cds	0.044
	107200	D20350	Hs.5628	ESTs	0.044
	101166	L14927	Hs.2099	lipocalin 1 (protein migrating faster than albumin; tear prealbumin)	0.044
	134289	M54915	Hs.81170	pim-1 oncogene	0.044
25	135329	AA436026	Hs.98858	ESTs	0.044
	124950	T03786	Hs.151531	protein phosphatase 3 (formerly 2B); catalytic subunit; beta isoform (calcineurin A beta)	0.044
	102919	X12447	Hs.183760	aldolase A; fructose-bisphosphate	0.044
	100574	HG2279-HT2375		Triosephosphate Isomerase	0.045
	131286	AA450092	Hs.25300	Homo sapiens clones 24718 and 24825 mRNA sequence	0.045
30	102675	U72512		Human B-cell receptor associated protein (hBAP) alternatively spliced mRNA, partial 3'UTR	0.045
	131332	R50487	Hs.25717	ESTs	0.045
	101634	M57731	Hs.75765	GRO2 oncogene	0.046
	113118	T47906	Hs.220512	ESTs	0.046
	124884	R77276	Hs.120911	ESTs	0.046
35	130523	W76097	Hs.214507	ESTs	0.046
	110244	H26742	Hs.25367	ESTs; Weakly similar to ALR [H.sapiens]	0.046
	131932	AA454980	Hs.25601	chromodomain helicase DNA binding protein 3	0.046
	132509	H09751	Hs.5038	neuropathy target esterase	0.046
	133372	AA291139	Hs.72242	ESTs	0.046
40	100817	HG4011-HT4804		Dystrophin-Associated Glycoprotein, 50 Kda, Alt. Splice 2	0.047
	106746	AA476436	Hs.7991	ESTs	0.047
	135401	L14813	Hs.169271	carboxyl ester lipase-like (bile salt-stimulated lipase-like)	0.047
	130479	R44163	Hs.12457	Homo sapiens clone 23770 mRNA sequence	0.047
	102589	U62015	Hs.8867	cysteine-rich; angiogenic inducer; 61	0.047
45	121521	AA412165	Hs.97358	EST	0.048
	135340	AA425137	Hs.99093	Homo sapiens chromosome 19; cosmid R28379	0.048
	132336	AA342422	Hs.45073	ESTs	0.048
	115368	AA282133	Hs.88960	ESTs; Weakly similar to similar to collagen [C.elegans]	0.048
	101278	L38487	Hs.110849	estrogen-related receptor alpha	0.048
50	103284	X80200	Hs.8375	TNF receptor-associated factor 4	0.048
	100564	HG2239-HT2324		Potassium Channel Protein.(Gb:Z11585)	0.048
	133132	Z40883	Hs.65588	ESTs; Weakly similar to dJ393P12.2 [H.sapiens]	0.048
	121811	AA424535	Hs.98416	ESTs	0.048
	129613	AA279481	Hs.238831	ESTs; Weakly similar to collagen alpha 1(XVIII) chain [M.musculus]	0.049
55	132468	S79854	Hs.49322	deiodinase; iodothyronine; type III	0.049
	120111	W95841	Hs.136031	ESTs	0.049
	103668	Z83741	Hs.248174	H2A histone family; member M	0.049
	130386	F10874	Hs.234249	mitogen-activated protein kinase 8 interacting protein 1	0.049
	104275	C02170	Hs.39387	ESTs; Weakly smlr to weak smlrity to ribosomal prot L14 [C.elegans]	0.049
60	106305	AA436146	Hs.12828	ESTs	0.05
	116431	AA609878	Hs.55289	ESTs; Weakly smlr to 110 KD CELL MEMBRANE GLYCOPROTEIN [H.sapiens]	0.813
	120339	AA206465	Hs.256470	EST	0.05
	114427	AA017063		ESTs; Highly similar to Miz-1 protein [H.sapiens]	0.05
	118821	N79070	Hs.94789	ESTs	0.05
65	118979	N93798	Hs.43666	protein tyrosine phosphatase type IVA; member 3	0.05
	107495	W78776	Hs.90375	ESTs	0.051
	120240	Z41732	Hs.66049	ESTs	0.051

5	114331	Z41309	Hs.12400	ESTs	0.051
	130947	R40037	Hs.21506	ESTs	0.052
	129242	W81679	Hs.5174	ribosomal protein S17	0.052
	131413	AA482390	Hs.26510	ESTs; Modly smlr to vacuolar prot sorting homolog r-vps33b [R.norvegicus]	0.052
	112304	R54798	Hs.26239	ESTs	0.052
10	101416	M17254	Hs.45514	v-ets avian erythroblastosis virus E26 oncogene related	0.052
	131201	AA426304	Hs.24174	ESTs	0.052
	101054	K02405	Hs.73933	Human MHC class II HLA-DQ-beta mRNA (DR7 DQw2); complete cds	0.052
	101306	L41143	Hs.232069	T-cell leukemia translocation altered gene	0.053
	129311	T55087		yb45c08.r1 Stratagene fetal spleen (#937205) Homo sapiens cDNA	
15	129942	U95301	Hs.144442	clone IMAGE:74126 5', mRNA sequence.	0.053
	119210	R93340	Hs.92995	phospholipase A2; group X	0.053
	101046	K01160		ESTs	0.053
	114086	Z38266	Hs.12770	Accession not listed in Genbank	0.053
	110171	H19964	Hs.31709	Homo sapiens PAC clone DJ0777O23 from 7p14-p15	0.053
20	101004	J04101	Hs.248109	ESTs	0.053
	129715	N58479	Hs.12126	v-ets avian erythroblastosis virus E26 oncogene homolog 1	0.053
	101581	M34996	Hs.198253	ESTs; Weakly similar to LR8 [H.sapiens]	0.053
	113285	T66830	Hs.182712	major histocompatibility complex; class II; DQ alpha 1	0.053
	127537	AA569531	Hs.162859	ESTs	0.053
25	100813	HG3995-HT4265		ESTs	0.054
	101841	M93107	Hs.76893	Cpg-Enriched Dna, Clone S19	0.054
	135053	R77159	Hs.93678	3-hydroxybutyrate dehydrogenase (heart; mitochondrial)	0.054
	101419	M17886	Hs.177592	ESTs	0.054
	119724	W69468	Hs.47622	ribosomal protein; large; P1	0.054
30	102673	U72509		ESTs	0.055
	129877	AA248589	Hs.13094	Human alternatively spliced B8 (B7) mRNA, partial sequence	0.055
	114788	AA156737	Hs.103904	ESTs; Weakly similar to ORF YGR101w [S.cerevisiae]	0.055
	123812	AA620607	Hs.111591	EST	0.055
	117669	N39237	Hs.44977	ESTs	0.055
35	123782	AA610111	Hs.162695	ESTs	0.055
	102395	U41767	Hs.92208	EST	0.055
	133795	M12529	Hs.169401	a disintegrin and metalloproteinase domain 15 (metargidin)	0.055
	123193	AA489228	Hs.136956	apolipoprotein E	0.055
	132595	AA253369	Hs.155742	ESTs	0.056
40	104161	AA456471	Hs.7724	glyoxylate reductase/hydroxypyruvate reductase	0.056
	115330	AA281145	Hs.88827	KIAA0963 protein	0.056
	112893	T08000	Hs.194684	ESTs	0.056
	133475	L29217	Hs.73987	bassoon (presynaptic cytomatrix protein)	0.056
	128699	K03207	Hs.103972	CDC-like kinase 3	0.056
45	102940	X13956	Hs.24998	proline-rich protein BstNI subfamily 4	0.056
	131299	AA431464	Hs.25426	Hu 12S RNA induced by poly(rI); poly(rC) and Newcastle disease virus	0.056
	102495	U51240	Hs.79356	ESTs; Weakly similar to unknown [H.sapiens]	0.057
	129594	R70379	Hs.115396	Lysosomal-associated multispinning membrane protein-5	0.057
	118593	N69020	Hs.207689	Human germline IgD chain gene; C-region; C-delta-1 domain	0.057
50	126702	U54602	Hs.2785	EST	0.057
	124386	N27368	Hs.212414	keratin 17	0.057
	130538	M20786	Hs.159509	sema domain; immunoglobulin domain (Ig); short basic domain;	0.057
	114299	Z40782	Hs.22920	secreted; (semaphorin) 3E	0.057
	115604	AA400378	Hs.49391	alpha-2-plasmin inhibitor	0.057
55	106052	AA416947	Hs.6382	similar to S68401 (cattle) glucose induced gene	0.057
	131730	U05681	Hs.31210	ESTs	0.057
	131285	AA479498	Hs.25274	ESTs; Highly similar to KIAA0612 protein [H.sapiens]	0.057
	129705	X78706	Hs.12068	B-cell CLL/lymphoma 3	0.057
	123175	AA489010	Hs.178400	ESTs; Modly smlr to putative seven pass transmembrane prot [H.sapiens]	0.058
60	103592	Z30644	Hs.123059	carnitine acetyltransferase	0.058
	118196	N59478	Hs.48396	ESTs	0.058
	104886	AA053348	Hs.144626	chloride channel Kb	0.058
	104250	AF000575	Hs.105928	ESTs; Moderately similar to tumor necrosis factor-alpha	0.058
				-induced protein B12 [H.sapiens]	0.058
65	113301	T67452	Hs.13104	growth differentiation factor 11	0.058
	110441	H50302	Hs.19845	leukocyte immunoglobulin-like receptor; subfamily B (with TM	0.058
	125297	Z39215	Hs.159409	and ITIM domains); member 3	0.058
	135258	AA292423	Hs.97272	EST	0.058
	130633	T92363	Hs.178703	ESTs; Highly smlr to prot phosphatase 2A BR gamma subunit [H.sapiens]	0.058
	112006	R42607	Hs.22241	ESTs	0.058
				hypothetical protein	0.058

5	130805	U12194	Hs.170238	sodium channel; voltage-gated; type I; beta polypeptide	0.058
	134907	D80002	Hs.178292	KIAA0180 protein	0.058
	132619	AA404565	Hs.53447	ESTs; Moderately similar to kinesin light chain 1 [M.musculus]	0.058
	135115	N35489	Hs.94653	neurochondrin	0.058
	100531	HG1872-HT1907		Major Histocompatibility Complex, Dg	0.058
10	124530	N62256	Hs.102727	EST	0.058
	119960	W87533	Hs.32699	ESTs; Moderately similar to LIV-1 protein [H.sapiens]	0.058
	132793	AA478999	Hs.56966	KIAA0906 protein	0.058
	101076	L04270	Hs.1116	lymphotoxin beta receptor (TNFR superfamily; member 3	0.058
	130655	N92934	Hs.17409	cysteine-rich protein 1 (intestinal)	0.058
15	134458	AA192614	Hs.83577	cysteine and glycine-rich protein 3 (cardiac LIM protein)	0.058
	105904	AA401452	Hs.32060	ESTs	0.059
	132878	AA026793	Hs.58679	ESTs; Weakly similar to 4F2/CD98 light chain [M.musculus]	0.059
	121828	AA425166	Hs.98497	ESTs	0.059
	133418	U76366	Hs.172727	Treacher Collins-Franceschetti syndrome 1	0.059
20	129317	N46244	Hs.110373	ESTs	0.059
	130153	D85815	Hs.15114	ras homolog gene family; member D	0.059
	124403	N31745	Hs.102493	ESTs	0.059
	127683	AA668123	Hs.134170	ESTs	0.059
	129814	W20070	Hs.168625	KIAA0979 protein	0.059
25	131770	D59682	Hs.31833	ESTs	0.06
	117557	N33920	Hs.44532	diubiquitin	0.06
	103522	Y10514		H.sapiens mRNA for CD152 protein	0.06
	120029	W91960	Hs.250640	sequence-specific single-stranded-DNA-binding protein	0.06
	102135	U15460	Hs.41691	activating transcription factor B	0.06
30	123617	AA609183	Hs.181131	ESTs	0.06
	112136	R46100	Hs.9739	ESTs	0.061
	133725	V00563	Hs.179543	immunoglobulin mu	0.061
	102069	U09196	Hs.82520	Hu 1.1 kb mRNA upregltd in retinoic acid treated HL-60 neutrophilic cells	0.061
	106555	AA455000	Hs.16725	ESTs	0.061
35	123269	AA491226	Hs.105280	ESTs; Weakly similar to dJ963K23.2 [H.sapiens]	0.061
	109088	AA166837	Hs.72620	DKFZP434I114 protein	0.061
	129399	AA263028	Hs.111076	malate dehydrogenase 2; NAD (mitochondrial)	0.061
	129375	W79850	Hs.11081	ESTs; Weakly similar to HPBRIL-7 protein [H.sapiens]	0.061
	135271	AA397763	Hs.97562	ESTs	0.061
40	132958	W90398	Hs.6147	KIAA1075 protein	0.061
	129364	AA477106	Hs.110757	DNA segment on chromosome 21 (unique) 2056 expressed sequence	0.061
	123427	AA598548	Hs.112471	ESTs	0.061
	105236	AA219179	Hs.19105	translocase of inner mitochondrial membrane 17 (yeast) homolog B	0.061
	101012	J04444	Hs.697	cytochrome c-1	0.062
45	134791	L18983	Hs.89655	protein tyrosine phosphatase; receptor type; N	0.062
	133700	K01396	Hs.75621	protease inhibitor 1 (anti-elastase); alpha-1-antitrypsin	0.062
	123887	AA621065	Hs.112943	ESTs	0.062
	129363	H05704	Hs.110746	H sapiens HCR (a-helix coiled-coil rod homologue) mRNA; complete cds	0.062
	105719	AA291644	Hs.36793	ESTs	0.062
50	124226	H62396	Hs.190266	ESTs	0.062
	117437	N27645		yw5e3.s1 Weizmann Olfactory Epithelium H sapiens cDNA clone	0.062
	132741	AA394133	Hs.55898	IMAGE:255676 3' smir to contains L1.13 L1 repetitive element ;, mRNA seq	0.062
	134437	M26041	Hs.198253	ESTs; Highly similar to OASIS protein [M.musculus]	0.062
	107664	AA010594	Hs.5326	major histocompatibility complex; class II; DQ alpha 1	0.062
55	120844	AA349417	Hs.96917	ESTs; Moderately similar to pim-1 protein [H.sapiens]	0.062
	101574	M34182	Hs.158029	ESTs	0.062
	131219	C00476	Hs.24395	protein kinase; cAMP-dependent; catalytic; gamma	0.062
	103495	Y09022	Hs.153591	small inducible cytokine subfamily B (Cys-X-Cys); member 14 (BRAK)	0.062
	129607	AA404594	Hs.11607	Not56 (D. melanogaster)-like protein	0.062
60	106467	AA450040	Hs.154162	ESTs	0.062
	128841	T16358	Hs.106443	ADP-ribosylation factor-like 2	0.062
	100515	HG1723-HT1729		ESTs	0.062
	119332	T54095		Macrophage Scavenger Receptor, Alt. Splice 2	0.062
	134516	AA171939	Hs.23413	ESTs; Weakly similar to !! ALU SUBFAMILY J WARNING ENTRY !! [H.sapiens]	0.062
65	135012	X73608	Hs.93029	ESTs	0.062
	103575	Z26256		sparc/osteonectin; cwcv and kazal-like domains proteoglycan (testican)	0.063
	115514	AA297739	Hs.55609	H.sapiens isoform 1 gene for L-type calcium channel, exon 1	0.063
	103996	AA321355		ESTs; Weakly similar to ISOLEUCYL-TRNA SYNTHETASE;	0.063
	110505	H55992	Hs.20495	CYTOPLASMIC [H.sapiens]	0.063
	133912	X62744	Hs.77522	EST2393 Bone marrow Homo sapiens cDNA 5' end, mRNA sequence	0.063
	129581	M33600	Hs.180255	DKFZP434F011 protein	0.063
				major histocompatibility complex; class II; DM alpha	0.063
				major histocompatibility complex; class II; DR beta 1	0.063

5	130139	R38280	Hs.150922	BCS1 (yeast homolog)-like	0.064
	105817	AA397825	Hs.5307	synaptopodin	0.064
	134658	AA410617	Hs.178009	ESTs	0.064
	100306	D50495	Hs.80598	transcription elongation factor A (SII); 2	0.064
	100277	D42053	Hs.75890	site-1 protease (subtilisin-like; sterol-regulated; cleaves sterol regulatory element binding proteins)	0.064
10	133116	D61259	Hs.6529	ESTs	0.064
	134909	AA521488	Hs.90998	KIAA0128 protein	0.064
	130319	X74794	Hs.154443	minichromosome maintenance deficient (S. cerevisiae) 4	0.064
	132057	AA102489	Hs.173484	ESTs	0.064
	108334	AA070473		zm7c8.s1 Stratagene neuroepithelium (#937231) Homo sapiens cDNA clone IMAGE:5399 3', mRNA sequence	0.064
15	129763	F10815	Hs.12373	KIAA0422 protein	0.064
	135112	T67464	Hs.94617	ESTs; Weakly similar to predicted using Genefinder [C.elegans]	0.064
	122269	AA436856	Hs.98910	ESTs	0.064
	133082	AA457129	Hs.6455	RuvB (E coli homolog)-like 2	0.064
	113213	T58607		ya94a02.s1 Stratagene placenta (#937225) Homo sapiens cDNA clone IMAGE:69290 3', mRNA sequence.	0.065
20	106228	AA429290	Hs.17719	ESTs	0.065
	130192	Y12661	Hs.171014	VGF nerve growth factor inducible	0.065
	104894	AA054087	Hs.18858	phospholipase A2; group IVC (cytosolic; calcium-independent)	0.065
	103508	Y10141		H.sapiens DAT1 gene, partial, VNTR	0.065
	128474	U40671	Hs.100299	ligase III; DNA; ATP-dependent	0.065
25	134012	AA417821	Hs.237924	ESTs; Highly similar to CGI-69 protein [H.sapiens]	0.065
	134536	AA457735	Hs.850	IMP (inosine monophosphate) dehydrogenase 1	0.065
	111714	R23146	Hs.23466	ESTs	0.065
	110521	H57060	Hs.108268	ESTs	0.065
	103282	X80198	Hs.77628	steroidogenic acute regulatory protein related	0.065
30	113921	W80730	Hs.28355	ESTs	0.065
	129331	N93465	Hs.110453	ESTs; Highly similar to CGI-38 protein [H.sapiens]	0.065
	111316	N74597	Hs.180535	ESTs; Weakly similar to mitogen inducible gene mig-2 [H.sapiens]	0.065
	135138	AA036794	Hs.95196	ESTs; Weakly similar to T20B12.3 [C.elegans]	0.065
	107289	T10792	Hs.172098	ESTs	0.065
35	121405	AA406083	Hs.98007	ESTs	0.065
	124965	T16275	Hs.106359	ESTs	0.065
	106595	AA456933	Hs.174481	ESTs	0.066
	100106	AF015910		Homo sapiens unknown protein mRNA, partial cds	0.066
	134715	AA282757	Hs.89040	prepronociceptin	0.066
40	135367	AA480109	Hs.9963	TYRO protein tyrosine kinase binding protein	0.066
	111533	R08548	Hs.251651	EST	0.066
	128509	R53109	Hs.247362	dimethylarginine dimethylaminohydrolase 2	0.066
	101030	J05037	Hs.76751	serine dehydratase	0.066
	102753	U80226		Human gamma-aminobutyric acid transaminase mRNA, partial cds	0.067
45	126991	R31652	Hs.821	biglycan	0.067
	109583	F02322	Hs.26135	ESTs	0.067
	119241	T12559	Hs.221382	ESTs	0.067
	130569	AA156597	Hs.256441	EST; Moderately similar to CGI-136 protein [H.sapiens]	0.067
	112926	T10316	Hs.4302	ESTs	0.067
50	120495	AA256073	Hs.190626	ESTs	0.067
	130931	AA278412	Hs.21346	ESTs; Weakly similar to F42C5.7 gene product [C.elegans]	0.067
	129982	M87789	Hs.140	immunoglobulin gamma 3 (Gm marker)	0.067
	133832	H03387	Hs.241305	estrogen-responsive B box protein	0.067
	110697	H93721	Hs.20798	ESTs	0.067
55	121183	AA400138	Hs.97703	ESTs	0.067
	130953	U12707	Hs.2157	Wiskott-Aldrich syndrome (eczema-thrombocytopenia)	0.067
	102218	U24183	Hs.75160	phosphofructokinase; muscle	0.067
	114181	Z39079	Hs.8021	KIAA1058 protein	0.067
	116581	D51287	Hs.82148	ribosomal protein S12	0.067
60	132498	T87708	Hs.50098	ESTs	0.068
	103788	AA096014	Hs.9527	ESTs; Highly similar to HSPC013 [H.sapiens]	0.068
	102459	U48936		Human amiloride-sensitive epithelial sodium channel gamma subunit mRNA, 5' end, partial cds	0.068
	100373	D79999	Hs.77225	ADP-ribosyltransferase (NAD+; poly (ADP-ribose) polymerase)-like 1	0.068
	132717	AA203321	Hs.151696	DKFZP727G051 protein	0.068
65	128863	D87462	Hs.106674	BRCA1 associated protein-1 (ubiquitin carboxy-terminal hydrolase)	0.068
	115193	AA262029	Hs.88218	ESTs	0.068
	124558	N66046	Hs.141605	ESTs	0.069
	117225	N20392	Hs.42846	ESTs	0.069
	110665	H83380	Hs.32757	ESTs	0.069

5	132905	U70663	Hs.182965	Kruppel-like factor 4 (gut)	0.069
	105778	AA348910	Hs.153299	DOM-3 (C. elegans) homolog Z	0.069
	134770	R72079	Hs.89575	CD79B antigen (immunoglobulin-associated beta)	0.069
	123097	AA485869	Hs.105671	ESTs	0.069
	100750	HG3523-HT4899		Proto-Oncogene C-Myc, Alt. Splice 3, Orf 114	0.069
10	125091	T91518		ye20f05.s1 Stratagene lung (#937210) H sapiens cDNA clone IMAGE: 3' similar to contains Alu repetitive element;contains MER12 repetitive element; mRNA sequence.	0.069
	100756	HG3565-HT3768		Zinc Finger Protein (Gb:M88357)	0.069
	113483	T87768	Hs.16439	ESTs	0.069
	101119	L09708	Hs.2253	complement component 2	0.069
	102286	U31628	Hs.12503	interleukin 15 receptor; alpha	0.07
15	135349	D83174	Hs.9930	collagen-binding protein 2 (colligen 2)	0.07
	100991	J03764	Hs.82085	plasminogen activator inhibitor; type 1	0.07
	133675	AA443720	Hs.7551	ESTs; Weakly similar to T25G3.1 [C.elegans]	0.07
	105422	AA251014	Hs.12210	ESTs	0.07
	102932	X13334	Hs.75627	CD14 antigen	0.07
20	119147	R58878	Hs.65739	ESTs	0.07
	104900	AA055048	Hs.180481	ESTs; Weakly similar to ACROSIN PRECURSOR [H.sapiens]	0.07
	133185	AA481404	Hs.6686	ESTs	0.07
	115496	AA290674	Hs.71819	eukaryotic translation initiation factor 4E binding protein 1	0.07
	121005	AA398332	Hs.97613	ESTs	0.07
25	124869	R69088	Hs.28728	ESTs; Weakly similar to F55A12.9 [C.elegans] ⁻	0.071
	129154	N23673	Hs.108969	mannosidase; alpha; class 2B; member 1	0.071
	112161	R48295		ESTs; Wkly smlr to !! ALU SUBFAMILY J WARNING ENTRY !! [H.sapiens]	0.071
	125251	W87486	Hs.141464	ESTs	0.071
	134298	J00116	Hs.81343	collagen; type II; alpha 1 (primary osteoarthritis; spondyloepiphyseal dysplasia; congenital)	0.071
30	119745	W70264	Hs.58093	ESTs	0.071
	131306	AA232686	Hs.25489	ESTs	0.071
	107776	AA018820	Hs.221147	ESTs	0.071
	134271	AA199630	Hs.184456	ESTs; Wkly smlr to !! ALU SUBFAMILY SX WARNING ENTRY !! [H.sapiens]	0.071
	101798	M85220		Accession not listed in Genbank	0.071
35	135402	S76942	Hs.99922	dopamine receptor D4	0.071
	118742	N74052	Hs.50424	EST	0.071
	131867	N64656	Hs.3353	Homo sapiens clone 24940 mRNA sequence	0.071
	102923	X12517	Hs.1063	small nuclear ribonucleoprotein polypeptide C	0.072
	100775	HG371-HT26388		Mucin 1, Epithelial, Alt. Splice 9	0.072
40	111020	N54361	Hs.185726	ESTs	0.072
	134224	X80822	Hs.163593	ribosomal protein L18a	0.072
	124059	F13673	Hs.99769	ESTs	0.072
	133972	AA160743	Hs.78019	Homo sapiens clone 24432 mRNA sequence	0.072
	129681	AA436009	Hs.178186	ESTs; Weakly similar to WASP-family protein [H.sapiens]	0.072
45	103065	X58399	Hs.81221	Human L2-9 transcript of unrearranged immunoglobulin V(H)5 pseudogene	0.072
	124966	T19271	Hs.155560	calnexin	0.072
	112270	R53021	Hs.203358	ESTs	0.072
	116704	F10183	Hs.66140	EST	0.072
	129890	M13699	Hs.111461	ceruloplasmin (ferroxidase)	0.072
50	127345	AA972008	Hs.166253	ESTs; Highly similar to KIAA0476 protein [H.sapiens]	0.072
	112436	R63090	Hs.28391	ESTs	0.072
	114531	AA053033	Hs.203330	ESTs	0.072
	135122	H99080	Hs.94814	ESTs	0.072
	103934	AA281338	Hs.134200	Homo sapiens mRNA; cDNA DKFZp564C186 (from clone DKFZp564C186)	0.072
55	109363	AA215369	Hs.185764	ESTs; Weakly similar to hypothetical protein [H.sapiens]	0.072
	112647	R83329	Hs.33403	ESTs	0.073
	127083	Z44079	Hs.91608	otoferlin	0.073
	133027	AA402624	Hs.63236	synuclein; gamma (breast cancer-specific protein 1)	0.073
	122086	AA432121	Hs.250986	EST	0.073
60	110405	H47542	Hs.33962	ESTs	0.073
	128697	AB002344	Hs.103915	KIAA0346 protein	0.073
	112221	R50380	Hs.25670	ESTs	0.073
	100478	HG1067-HT1067		Mucin (Gb:M22406)	0.073
	115598	AA400129	Hs.65735	ESTs	0.073
65	132491	AA227137	Hs.4984	KIAA0828 protein	0.073
	101655	M60299		Human alpha-1 collagen type II gene, exons 1, 2 and 3	0.073
	106018	AA411887	Hs.34737	ESTs	0.073
	129683	W05348	Hs.158196	DKFZP434B103 protein	0.073
	134137	F10045	Hs.79347	KIAA0211 gene product	0.073
	114008	W89128	Hs.19872	ESTs	0.073

	107653	AA010210	Hs.47041	ESTs	0.073
	104798	AA029462	Hs.17235	ESTs	0.073
	134082	L16991	Hs.79006	deoxythymidylate kinase	0.073
5	119180	R80413	Hs.92520	ESTs	0.073
	107741	AA016982	Hs.64341	ESTs	0.073
	133683	AA335223	Hs.75558	pepsinogen 5; group I (pepsinogen A)	0.073
	111694	R22035	Hs.23331	ESTs	0.073
	120764	AA338729	Hs.133096	ESTs	0.073
	119389	T88826	Hs.90973	ESTs	0.074
10	100929	HG688-HT688		Major Histocompatibility Complex, Class II, Dr Beta 2 (Gb:X65561)	0.074
	119388	T88798		plasminogen activator inhibitor; type I	0.074
	133019	AF009674	Hs.184434	axin	0.074
	105185	AA191495	Hs.189937	ESTs	0.074
	133413	S72043	Hs.73133	metallothionein 3 (growth inhibitory factor (neurotrophic))	0.074
15	101017	J04599	Hs.821	biglycan	0.074
	132865	K02765	Hs.251972	complement component 3	0.074
	110882	N36001	Hs.17348	ESTs; Wkly smlr to !! ALU SUBFAMILY SQ WARNING ENTRY !! [H.sapiens]	0.074
	129197	T90303	Hs.109308	ESTs; Wkly smlr to leucine-rich glioma-inactivated prot precursor [H.sapiens]	0.074
	101184	L19871	Hs.460	activating transcription factor 3	0.075
20	134910	AA431320	Hs.9100	ESTs	0.075
	119411	T96621	Hs.203656	EST	0.075
	102000	U01824	Hs.380	solute carrier family 1 (glial high affinity glutamate transporter); member 2	0.075
	114691	AA121893	Hs.103779	ESTs; Weakly similar to envelope protein [H.sapiens]	0.075
	134179	U53204	Hs.79706	plectin 1; intermediate filament binding protein; 500kD	0.075
25	134503	U34880	Hs.84183	diphtheria toxin resistance protein required for diphthamide biosynthesis (Saccharomyces)-like 1	0.075
	129719	N66396	Hs.167766	ESTs; Moderately similar to Pro-a2(XI) [H.sapiens]	0.075
	113916	W80464	Hs.31928	ESTs; Wkly smlr to alternatively spliced product using exon 13A [H.sapiens]	0.075
	113897	W73926	Hs.4947	ESTs	0.075
30	129697	R00841	Hs.172069	DKFZP434C212 protein	0.075
	112078	R44155	Hs.112218	ESTs	0.075
	121980	AA429886	Hs.110407	ESTs; Weakly similar to coded for by C. elegans cDNA yk173c12.5 [C.elegans]	0.075
	100898	HG4638-HT5050		Spliceosomal Protein Sap 49	0.075
	121626	AA416974	Hs.98174	ESTs	0.075
35	133670	AA243416	Hs.75470	hypothetical protein; expressed in osteoblast	0.075
	131879	AA017161	Hs.33792	ESTs	0.075
	100254	D38037	Hs.77643	FK506-binding protein 1B (12.6 kD)	0.075
	133194	AA291726	Hs.67201	ESTs	0.075
	106081	AA418394	Hs.25354	ESTs	0.075
40	115544	AA351433	Hs.66187	Homo sapiens clone 23700 mRNA sequence	0.076
	119955	W87460	Hs.58989	ESTs	0.076
	104407	H61361	Hs.102171	immunoglobulin superfamily containing leucine-rich repeat	0.076
	135019	X58431	Hs.98428	Human Hox2.2 gene for a homeobox protein	0.076
	114815	AA161488	Hs.103931	DKFZP434B0335 protein	0.076
45	119471	W31352	Hs.55445	ESTs	0.076
	117788	N48292	Hs.46849	ESTs	0.076
	119406	T95064	Hs.193771	EST	0.076
	130777	R61742	Hs.256554	ESTs	0.076
	130494	L13197	Hs.75874	pregnancy-associated plasma protein A	0.076
50	104107	AA424111	Hs.12598	T-cell lymphoma invasion and metastasis 2	0.076
	121483	AA411981	Hs.25274	ESTs; Modly smlr to putative seven pass transmembrane prot [H.sapiens]	0.076
	104451	M13299	Hs.102119	blue cone pigment	0.076
	118027	N52770	Hs.75968	thymosin; beta 4; X chromosome	0.076
	109419	AA227560	Hs.86987	receptor-interacting serine-threonine kinase 3	0.076
55	115783	AA424487	Hs.72289	ESTs; Weakly similar to LIV-1 protein [H.sapiens]	0.076
	110585	H62223	Hs.133526	ESTs; Wkly smlr to !!!ALU SUBFAMILY SB1 WARNING ENTRY !!! [H.sapiens]	0.076
	123165	AA488863	Hs.105216	ESTs; Weakly smlr to !!!ALU SUBFAMILY J WARNING ENTRY !!! [H.sapiens]	0.077
	103966	AA303166	Hs.127270	ESTs	0.077
	109549	F01528	Hs.21192	Homo sapiens clone 25155 mRNA sequence	0.077
60	106730	AA465520	Hs.22313	ESTs	0.077
	120310	AA193676	Hs.118926	DKFZP586K0919 protein	0.077
	104078	AA402801	Hs.222010	ESTs	0.077
	117624	N35978	Hs.82364	ESTs	0.077
	112421	R62441	Hs.23127	ESTs	0.077
65	106958	AA497026	Hs.22059	ESTs	0.077
	129984	W92811	Hs.183927	ESTs; Weakly similar to !! ALU SUBFAMILY J WARNING ENTRY !! [H.sapiens]	0.077
	122044	AA431456	Hs.98736	EST	0.077
	123280	AA491285	Hs.175144	ESTs	0.077
	115710	AA412535	Hs.55235	sphingomyelin phosphodiesterase 2; neutra	0.077

				I membrane (neutral sphingomyelinase)	0.077
				KIAA0255 gene product	0.077
				Homo sapiens clone 643 unknown mRNA; complete sequence	0.078
5	130513	AA460257	Hs.15866	ESTs	0.078
	100996	J03909	Hs.14623	interferon; gamma-inducible protein 30	0.078
	128358	AI095718	Hs.135015	ESTs	0.078
	128544	R59352	Hs.119273	KIAA0296 gene product	0.078
	106040	AA412681	Hs.125139	ESTs	0.078
10	106495	AA452113	Hs.32454	ESTs; Moderately similar to KIAA0544 protein [H.sapiens]	0.078
	131833	R40899	Hs.32973	glycine receptor; beta	0.078
	119219	R97176	Hs.110783	ESTs	0.078
	135415	X60655	Hs.99967	even-skipped homeo box 1 (homolog of Drosophila)	0.078
	109457	AA232646	Hs.68061	ESTs; Weakly similar to sphingosine kinase [M.musculus]	0.078
	117137	H96670	Hs.42221	ESTs	0.078
15	107094	AA609614	Hs.5241	ESTs	0.078
	130165	T90529	Hs.251613	EST	0.078
	124072	H05252	Hs.101637	EST; Weakly similar to hypothetical protein [H.sapiens]	0.078
	126151	AA324743	Hs.40808	ESTs	0.078
	119035	R01779	Hs.7740	ESTs	0.078
20	110157	H18997	Hs.169731	ESTs	0.078
	128515	AA149044	Hs.10086	ESTs; Highly similar to HYPOTHETICAL PROTEIN KIAA0195 [H.sapiens]	0.078
	133069	U94836	Hs.6430	protein with polyglutamine repeat	0.078
	112209	R49644	Hs.24865	ESTs	0.078
	133361	R28279	Hs.71848	Human clone 23548 mRNA sequence	0.078
25	134714	U89922	Hs.890	lymphotoxin beta (TNF superfamily; member 3)	0.078
	129905	T86796	Hs.132875	ESTs; Weakly similar to predicted using Genefinder [C.elegans]	0.079
	120421	AA236166	Hs.132957	ESTs; Weakly similar to chondromodulin-I precursor [H.sapiens]	0.079
	100885	HG4490-HT4876		Proline-Rich Protein Prb4, Allele	0.079
	102789	U86759	Hs.158336	netrin 2 (chicken)-like	0.079
30	120139	Z39273	Hs.77876	Human DNA from chromosome 19-specific cosmid R30923; genomic sequence	0.079
	135238	U76343	Hs.96970	Human liver GABA transport protein mRNA; 3' end	0.079
	129618	N54845	Hs.173030	ESTs	0.079
	132960	AA609742	Hs.6150	KIAA0521 protein	0.079
	108751	AA127063	Hs.203717	ESTs	0.079
35	134060	D42039	Hs.78871	KIAA0081 protein	0.079
	111338	N79778	Hs.35094	extracellular matrix protein 2; female organ and adipocyte specific	0.079
	112345	R56880	Hs.26563	ESTs	0.079
	126456	W00881		za56d02.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone	0.079
40	128937	Z39939	Hs.10726	IMAGE:296547 5', mRNA sequence.	0.079
	103485	Y08409	Hs.248415	ESTs	0.079
	111202	N68280	Hs.107922	thyroid hormone responsive SPOT14 (rat) homolog	0.079
	132625	AA429890	Hs.166066	ESTs	0.079
	103434	X98085	Hs.54433	cisplatin resistance associated	0.079
45	102616	U65581	Hs.159191	tenascin R (restrictin; janusin)	0.079
	102667	U70867	Hs.83974	ribosomal protein L3-like	0.079
	111422	R01127	Hs.19104	solute carrier family 21 (prostaglandin transporter); member 2	0.079
	101411	M16938	Hs.820	ESTs	0.079
	113267	T65058	Hs.12725	homeo box C6	0.08
50	103559	Z19585	Hs.75774	ESTs; Weakly similar to !! ALU SUBFAMILY J WARNING ENTRY !! [H.sapiens]	0.08
	131588	AA258613	Hs.29189	thrombospondin 4	0.08
	107821	AA020991	Hs.172856	KIAA1021 protein	0.08
	134278	H82839	Hs.81001	ESTs	0.08
	120893	AA369800	Hs.97058	ESTs; Weakly similar to DY3.6 [C.elegans]	0.08
55	108786	AA128999		EST; Highly similar to CMP-N-acetylneuraminic acid hydroxylase [H.sapiens]	0.08
	106890	AA489245	Hs.88500	zo8f12.s1 Stratagene neuroepithelium NT2RAMI 937234 Homo sapiens	0.08
	119760	W72267	Hs.58219	cDNA clone IMAGE:567119 3', mRNA sequence	0.08
	132999	Y00787	Hs.624	KIAA1066 protein; JSAP1 homolog (mouse); JIP3 homolog (mouse)	0.08
60	129156	AA028195	Hs.108973	ESTs	0.08
	121171	AA400008	Hs.161814	interleukin 8	0.08
	103864	AA207264	Hs.181077	dolichyl-phosphate mannosyltransferase polypeptide 2; regulatory subunit	0.08
	128591	AA255537	Hs.102057	ESTs	0.08
	122172	AA435753	Hs.161854	ESTs; Weakly similar to Miller-Dieker lissencephaly gene [H.sapiens]	0.08
65	112802	R97647	Hs.174855	ESTs; Weakly similar to O-linked GlcNAc transferase [H.sapiens]	0.08
	107723	AA015967	Hs.60680	EST	0.08
	113011	T23737	Hs.1600	EST	0.08
	131279	AA089853	Hs.25197	chaperonin containing TCP1; subunit 5 (epsilon)	0.081
	103190	X70083	Hs.58414	STIP1 homology and U-Box containing protein 1	0.081
				filamin C; gamma (actin-binding protein-280)	0.081

	103956	AA292411	Hs.233348	ESTs	0.081
	112706	R89828	Hs.138493	ESTs	0.081
	126126	M85370		EST01884 Fetal brain, Stratagene (cat#936206) Homo sapiens cDNA clone HFBCH10, mRNA sequence.	0.081
5	130094	H43286	Hs.167017	gamma-aminobutyric acid (GABA) B receptor; 1	0.081
	100800	HG3945-HT4215		Phospholipid Transfer Protein	0.081
	108675	AA115240	Hs.61816	ESTs	0.081
	129420	AA234259	Hs.99816	ESTs	0.081
10	129666	M77349	Hs.118787	transforming growth factor; beta-induced; 68kD	0.081
	101645	M59807	Hs.943	natural killer cell transcript 4	0.081
	130536	T17045	Hs.159492	spastic ataxia of Charlevoix-Saguenay (sacsin)	0.081
	107732	AA016181	Hs.59752	ESTs	0.081
	123071	AA482593	Hs.104285	ESTs	0.081
	113537	T90457	Hs.191293	ESTs	0.081
15	101250	L34060	Hs.79133	cadherin 8	0.081
	122521	AA449433	Hs.149227	ESTs; Weakly similar to PROLINE-RICH PROTEIN MP-3 [M.musculus]	0.081
	133914	N32811	Hs.77542	ESTs	0.081
	102038	U05659	Hs.477	hydroxysteroid (17-beta) dehydrogenase 3	0.081
	110336	H40338	Hs.174094	ESTs; Weakly similar to !! ALU SUBFAMILY J WARNING ENTRY !! [H.sapiens]	0.081
20	118637	N70274	Hs.49822	ESTs	0.081
	117966	N51589	Hs.94012	ESTs	0.082
	104424	H87671	Hs.182320	ESTs; Weakly similar to Mouse 19.5 mRNA; complete cds [M.musculus]	0.082
	100361	D78361	Hs.125078	Human mRNA for ornithine decarboxylase antizyme; ORF 1 and ORF 2	0.082
	112974	T17291	Hs.101174	microtubule-associated protein tau	0.082
25	123832	D63482	Hs.57734	KIAA0148 gene product	0.082
	132039	Z39489	Hs.3781	Homo sapiens BAC clone RG118D07 from 7q31	0.082
	113272	T65383	Hs.12807	ESTs	0.082
	104924	AA058532	Hs.28774	ESTs	0.082
	111061	N58054	Hs.36859	ESTs	0.082
30	129269	R45977	Hs.163593	ribosomal protein L18a	0.082
	102453	U48437	Hs.74565	amyloid beta (A4) precursor-like protein 1	0.082
	126204	AI080388	Hs.134296	ESTs	0.082
	116615	D80666	Hs.45203	ESTs	0.082
	128856	AA219552	Hs.204144	ESTs; Modly smlr to tumor necrosis factor-alpha-induced prot B12 [H.sapiens]	0.082
35	112776	R95850	Hs.34494	ESTs	0.082
	105494	AA256273	Hs.29288	Homo sapiens mRNA; cDNA DKFZp434P174 (from clone DKFZp434P174)	0.082
	117000	H84718	Hs.112236	ESTs; Weakly similar to repressor protein [H.sapiens]	0.082
	112656	R85260	Hs.133151	transient receptor potential channel 7	0.082
40	128963	J03890	Hs.1074	surfactant; pulmonary-associated protein C	0.083
	116957	H79292	Hs.39960	ESTs	0.083
	101057	K03430		Human complement C1q B-chain gene, exon A+1	0.083
	121948	AA429452	Hs.98582	ESTs	0.083
	130822	M80647	Hs.2001	thromboxane A synthase 1 (platelet; cytochrome P450; subfamily V)	0.083
45	122743	AA458674	Hs.99478	EST	0.083
	114569	AA063316		zm2d1.s1 Stratagene corneal stroma (#937222) Homo sapiens cDNA clone IMAGE:512947 3' similar to TR:E198281 E198281 THIOREDOXIN REDUCTASE ;contains Alu repetitive element; mRNA sequence	0.083
	132270	U70671	Hs.43509	ataxin 2 related protein	0.083
50	108126	AA052951	Hs.47413	ESTs	0.083
	102880	X04325	Hs.2679	gap junction protein; beta 1; 32kD (connexin 32; Charcot-Marie-Tooth neuropathy; X-linked)	0.083
	115365	AA282089	Hs.88599	ESTs	0.083
	114529	AA052980	Hs.206704	ESTs	0.083
55	135017	AA249586	Hs.9315	ESTs; Weakly similar to NEURONAL OLFACTOMEDIN-RELATED ER LOCALIZED PROTEIN [H.sapiens]	0.083
	123776	AA610071	Hs.112813	ESTs	0.083
	114454	AA021091	Hs.226208	ESTs	0.083
	101246	L33799	Hs.202097	procollagen C-endopeptidase enhancer	0.083
60	107366	U78310	Hs.13501	pescadillo (zebrafish) homolog 1; containing BRCT domain	0.083
	132779	T89601	Hs.95497	ESTs; Weakly similar to GLUCOSE TRANSPORTER TYPE 5; SMALL INTESTINE [H.sapiens]	0.083
	129709	AA112209	Hs.1209	acyl-Coenzyme A dehydrogenase; long chain	0.083
65	115244	AA278767	Hs.914	Human mRNA for SB classII histocompatibility antigen alpha-chain	0.083
	123253	AA490878	Hs.111334	ferritin; light polypeptide	0.083
	128469	T23724	Hs.258677	EST	0.083
	132220	AA431847	Hs.42409	ESTs; Highly similar to CGI-146 protein [H.sapiens]	0.083
	111664	R17939	Hs.22344	ESTs	0.083
	102354	U38268		Human cytochrome b pseudogene, partial cds	0.084
	112828	R98774	Hs.194338	ESTs	0.084

5	110410	H47868	Hs.34024	ESTs	0.084
	102620	U66052		Human clone W2-6 mRNA from chromosome X	0.084
	102550	U58087	Hs.14541	cullin 1	0.084
	108417	AA075716		zm89a5.s1 Stratagene ovarian cancer (#937219) H sapiens cDNA clone IMAGE:54512 3' similar to gb:X14723 CLUSTERIN PRECURSOR (HUMAN);, mRNA sequence	0.084
10	113299	T67285	Hs.13089	ESTs	0.084
	117869	N49947	Hs.46990	ESTs	0.084
	113734	T98484	Hs.18377	EST	0.084
	133325	C00424	Hs.7101	periodontal ligament fibroblast protein	0.084
15	123368	AA505022	Hs.124838	ESTs	0.084
	101615	M55153	Hs.8265	transglutaminase 2 (C polypeptide; protein-glutamine -gamma-glutamyltransferase)	0.084
	119352	T65972	Hs.193365	ESTs; Moderately similar to alternatively spliced product using exon 13A [H.sapiens]	0.084
	123828	AA620686	Hs.112884	EST	0.084
20	103611	Z38133	Hs.113973	myosin; heavy polypeptide 8; skeletal muscle; perinatal	0.084
	131289	AA485697	Hs.25334	ESTs; Weakly similar to ION CHANNEL HOMOLOG RIC PRECURSOR [M.musculus]	0.084
	128678	T15896	Hs.103535	ESTs	0.084
	130814	AA256695	Hs.19813	ESTs	0.084
25	133391	X57579	Hs.727	inhibin; beta A (activin A; activin AB alpha polypeptide)	0.084
	129322	AA437153	Hs.110407	ESTs; Weakly similar to coded for by C. elegans cDNA yk173c12.5 [C.elegans]	0.084
	109284	AA196995	Hs.86092	ESTs	0.085
	116689	F09222	Hs.66099	ESTs	0.085
30	100545	HG2147-HT2217		Mucin 3, Intestinal (Gb:M55405)	0.085
	102634	U66711	Hs.77667	lymphocyte antigen 6 complex; locus E	0.085
	111735	R25389	Hs.23856	ESTs; Weakly similar to FAST kinase [H.sapiens]	0.085
	105181	AA190676	Hs.10974	ESTs; Moderately similar to unknown [R.norvegicus]	0.085
35	122681	AA455350	Hs.99401	EST	0.085
	114543	AA056121	Hs.158419	ESTs	0.085
	133597	AA425908	Hs.75139	partner of RAC1 (arfaptin 2)	0.085
	121064	AA398647	Hs.97406	ESTs	0.085
40	122231	AA436369	Hs.197728	ESTs; Weakly similar to ZINC FINGER PROTEIN 135 [H.sapiens]	0.085
	100309	D50550	Hs.95659	lethal giant larvae (Drosophila) homolog 1	0.085
	101727	M73481	Hs.73883	gastrin-releasing peptide receptor	0.085
	131226	AA165400	Hs.24476	ESTs	0.085
45	133580	AA095041	Hs.181073	ESTs	0.085
	102792	U87964	Hs.227576	GTP binding protein 1	0.085
	104976	AA086480	Hs.183669	ESTs; Weakly similar to !! ALU SUBFAMILY J WARNING ENTRY !! [H.sapiens]	0.085
	120865	AA350631	Hs.96963	EST	0.085
50	106080	AA418046	Hs.35124	ESTs	0.085
	128571	AA416619	Hs.101661	ESTs	0.085
	101838	M92934	Hs.75511	connective tissue growth factor	0.085
	128514	H84261	Hs.100843	ESTs; Weakly similar to similar to GTP-binding protein [C.elegans]	0.085
55	123099	AA485931	Hs.79	aminoacylase 1	0.085
	134067	Y08200	Hs.78920	Rab geranylgeranyltransferase; alpha subunit	0.085
	116967	H80336	Hs.40124	EST	0.085
	110053	H12586	Hs.89563	nuclear cap binding protein 1; 80kD	0.085
60	114395	AA007313	Hs.110155	ESTs	0.085
	107465	W44681	Hs.251385	murine retrovirus integration site 1 homolog	0.085
	101983	S85655	Hs.75323	prohibitin	0.085
	112544	R70948	Hs.29153	ESTs	0.086
65	111423	R01165	Hs.188507	ESTs	0.086
	127918	AA806043	Hs.115396	Human germline IgD chain gene; C-region; C-delta-1 domain	0.086
	107300	T40348	Hs.90488	ESTs	0.086
	134947	R51194		yj71a08.r1 Soares breast 2NbHBst Homo sapiens cDNA clone IMAGE:154166 5' similar to gb:L11284 DUAL SPECIFICITY MITOGEN-ACTIVATED PROTEIN KINASE KINASE 1 (HUMAN);, mRNA sequence.	0.086
70	124579	N68345	Hs.127179	ESTs; Weakly similar to TERATOCARCINOMA-DERIVED GROWTH FACTOR 1 [H.sapiens]	0.086
	130471	Z68280	Hs.183706	adducin 1 (alpha)	0.086
	116596	D60755	Hs.92955	ESTs	0.086
	105069	AA136345	Hs.23617	ESTs; Weakly similar to ZFOC1 gene product [H.sapiens]	0.086
75	102491	U51010		Human nicotinamide N-methyltransferase gene, exon 1 and 5' flanking region	0.086
	130069	AA055896	Hs.146428	collagen; type V; alpha 1	0.086
	130234	AA280413	Hs.157441	spleen focus forming virus (SFFV) proviral integration oncogene spi1	0.086
	120540	AA262992	Hs.96417	ESTs	0.086
80	122508	AA449221	Hs.20432	ESTs	0.086

5	128054	AI205718	Hs.125416	ESTs	0.086
	133020	AA053248	Hs.185182	ESTs; Highly similar to 40S RIBOSOMAL PROTEIN S10 [H.sapiens]	0.086
	130056	AA017356	Hs.171900	armadillo repeat gene deletes in velocardiofacial syndrome	0.086
	130504	U48865	Hs.158323	CCAAT/enhancer binding protein (C/EBP); epsilon	0.086
	133978	W73859	Hs.78061	transcription factor 21	0.086
10	105265	AA227941	Hs.26088	ESTs	0.086
	133035	T15965	Hs.6333	ESTs	0.086
	100768	HG3636-HT3846		Myosin, Heavy Polypeptide 9, Non-Muscle	0.086
	129338	T56800	Hs.47274	Homo sapiens mRNA; cDNA DKFZp564B176 (from clone DKFZp564B176)	0.086
	132789	W23761	Hs.56876	ESTs	0.086
15	116099	AA456309	Hs.58831	regulator of Fas-induced apoptosis	0.086
	100721	HG3355-HT3532		Peroxisome Proliferator Activated Receptor (Gb:Z30972)	0.087
	112569	R73150	Hs.75270	GTP-binding protein homologous to Saccharomyces cerevisiae SEC4	0.087
	130645	AA020942	Hs.17200	STAM-like protein containing SH3 and ITAM domains 2	0.087
	100751	HG3527-HT3721		Luteinizing Hormone, Beta Subunit	0.087
20	134550	M27161	Hs.85258	CD8 antigen; alpha polypeptide (p32)	0.087
	130885	AA338646	Hs.20912	adenomatous polyposis coli like	0.087
	101446	M21302	Hs.56306	small proline-rich protein 2A	0.087
	116287	AA487856	Hs.155829	KIAA0676 protein	0.087
	134034	X89267	Hs.78601	uroporphyrinogen decarboxylase	0.087
25	130860	U68061	Hs.241395	protease; serine; 1 (trypsin 1)	0.087
	109901	H04992	Hs.30499	ESTs	0.087
	107537	Z20777	Hs.9857	ESTs; Weakly similar to peroxisomal short-chain alcohol dehydrogenase [H.sapiens]	0.087
	133232	AA496030	Hs.6845	ESTs	0.087
	108559	AA085161		zn12c5.s1 Stratagene hNT neuron (#937233) H sapiens cDNA clone IMAGE:54728 3' similar to TR:G1151228 G1151228 LPG1P. ; mRNA seq	0.087
30	121288	AA401735	Hs.97340	EST	0.087
	108844	AA132916	Hs.177961	Human Chromosome 16 BAC clone CIT987SK-A-388D4	0.087
	129874	AA406488	Hs.181551	ESTs	0.087
	105139	AA164543	Hs.110082	ESTs	0.088
	124789	R43803	Hs.78110	ESTs; Weakly similar to F17A9.2 [C.elegans]	0.088
35	115923	AA441929	Hs.38205	ESTs	0.088
	123640	AA609292	Hs.112681	ESTs	0.088
	131607	AA351409	Hs.172740	microtubule-associated protein; RP/EB family; member 3	0.088
	130064	T67053	Hs.181125	immunoglobulin lambda gene cluster	0.088
	108752	AA127070	Hs.71055	ESTs	0.088
40	124249	H68077	Hs.108211	ESTs	0.088
	100109	AJ000480	Hs.143513	phosphoprotein regulated by mitogenic pathways	0.088
	104642	AA004662	Hs.184245	KIAA0929 protein Mx2 interacting nuclear target (MINT) homolog	0.088
	131752	AA453311	Hs.31566	ESTs	0.088
	114727	AA132545	Hs.190202	ESTs	0.088
45	120965	AA398089	Hs.179715	ESTs	0.088
	100396	D84361	Hs.151123	Human mRNA for p52 and p64 isoforms of N-Shc; complete cds	0.088
	106218	AA428451	Hs.91146	DKFZP586E0820 protein	0.088
	111562	R09567	Hs.187569	ESTs	0.088
	121219	AA400606	Hs.144344	EST	0.088
50	101187	L20316	Hs.208	glucagon receptor	0.088
	101513	M28210	Hs.27744	RAB3A; member RAS oncogene family	0.088
	116454	AA621071	Hs.42034	ESTs; Moderately similar to T-complex protein 10A [H.sapiens]	0.088
	116171	AA463434	Hs.42658	ESTs	0.089
	117500	N31909	Hs.44278	ESTs	0.089
55	119978	W88623	Hs.59190	EST	0.089
	132005	D58231	Hs.173091	DKFZP434K151 protein	0.089
	109914	H05529	Hs.194704	leucine-rich; glioma inactivated 1	0.089
	130370	M55265	Hs.155140	casein kinase 2; alpha 1 polypeptide	0.089
	104262	AF009801	Hs.105941	bagpipe homeobox (Drosophila) homolog 1	0.089
60	129708	AA417181	Hs.120858	ESTs	0.089
	106398	AA447545	Hs.18268	adenylate kinase 5	0.089
	120884	AA365356	Hs.97041	ESTs	0.089
	130404	X72012	Hs.76753	endoglin (Osler-Rendu-Weber syndrome 1)	0.089
	114072	Z38184	Hs.123633	ESTs	0.089
65	131470	X54938	Hs.2722	inositol 1,4;5-trisphosphate 3-kinase A	0.089
	124573	N67935	Hs.194703	adaptor-related protein complex 4; mu 1 subunit	0.089
	114717	AA131240	Hs.252014	EST	0.089
	133806	M12759	Hs.76325	Human Ig J chain gene	0.09
	130470	AA398552	Hs.15711	KIAA0639 protein	0.09
	133182	Z80787	Hs.240135	H4 histone family; member J	0.09
	116036	AA452572	Hs.43866	ESTs	0.09

5	132404	AA393903	Hs.4768	ESTs	0.09
	122695	AA456048	Hs.99403	ESTs; Moderately similar to undulin 2 [H.sapiens]	0.09
	125975	AA495891	Hs.152290	ESTs; Highly similar to PACAP type-3/VIP type-2 receptor [H.sapiens]	0.09
	110783	N23669	Hs.26407	ESTs	0.09
	129860	AA410343	Hs.129826	tetraspan transmembrane 4 super family	0.09
10	120740	AA302650	Hs.96654	EST	0.09
	119564	W38206		Accession not listed in Genbank	0.09
	134474	AA054746	Hs.8379	ESTs	0.09
	119014	N95435	Hs.55144	ESTs	0.09
	109791	F10669	Hs.13228	DRE-antagonist modulator; calsenilin	0.09
15	117605	N35073	Hs.44433	ESTs	0.09
	121589	AA416627	Hs.191598	ESTs	0.09
	104326	D81655	Hs.143067	ESTs	0.09
	129861	N69507	Hs.129849	DKFZP564M182 protein	0.09
	102795	U88667	Hs.198396	ATP-binding cassette; sub-family A (ABC1); member 4	0.09
20	119626	W49499	Hs.184456	ESTs; Wkly smlr to !! ALU SUBFAMILY SX WARNING ENTRY !! [H.sapiens]	0.09
	110516	H56894	Hs.37368	EST	0.09
	105382	AA236853	Hs.111801	Homo sapiens mRNA; cDNA DKFZp564H2023 (from clone DKFZp564H2023)	0.09
	123754	AA609964	Hs.102021	ESTs	0.09
	108008	AA039430	Hs.61920	ESTs	0.09
25	121057	AA398619	Hs.142375	ESTs; Moderately similar to putative envelope protein [H.sapiens]	0.091
	123675	AA609474	Hs.112713	EST	0.091
	135194	C20975	Hs.9613	ESTs; Highly similar to angiopoietin-related protein [H.sapiens]	0.091
	127070	AA641812	Hs.190037	ESTs	0.091
	134051	S67070	Hs.78846	heat shock 27kD protein 2	0.091
30	133382	AA112532	Hs.7247	ESTs	0.091
	103615	Z46967	Hs.115460	calicin	0.091
	118457	N66593	Hs.49230	EST	0.091
	118504	N67334	Hs.50158	ESTs	0.091
	112915	T10176	Hs.4254	ESTs	0.091
35	132088	AA470121	Hs.243960	HLA-B associated transcript-3	0.091
	101504	M27288	Hs.248156	oncostatin M	0.091
	112550	R71391	Hs.29074	ESTs	0.091
	128551	H09058	Hs.237323	N-acetylglucosamine-phosphate mutase; DKFZP434B187 protein	0.091
	112879	T03541	Hs.115960	ESTs	0.091
40	127079	AI364691	Hs.128628	ESTs; Moderately similar to CL3BC [R.norvegicus]	0.091
	101993	U01062	Hs.77515	inositol 1,4,5-triphosphate receptor; type 3	0.091
	113020	T23830	Hs.7303	ESTs; Weakly similar to PROHIBITIN [H.sapiens]	0.091
	120465	AA251505	Hs.130861	ESTs	0.091
	130152	U32645	Hs.151139	E74-like factor 4 (ets domain transcription factor)	0.091
45	104941	AA065169	Hs.17805	ESTs	0.091
	110090	H16076	Hs.6915	ESTs	0.091
	135375	AA480888	Hs.99741	ESTs; Weakly similar to BRAIN PROTEIN H5 [H.sapiens]	0.091
	123799	AA620418	Hs.112861	ESTs	0.092
	118966	N93438	Hs.76907	ESTs; Highly similar to HSPC002 [H.sapiens]	0.092
50	116969	H80633	Hs.143038	ESTs	0.092
	125147	W38150		Accession not listed in Genbank	0.092
	100836	HG4113-HT4383		Olfactory Receptor Or17-201	0.092
	114726	AA132509	Hs.103827	EST	0.092
	107311	T57738	Hs.174112	ESTs	0.092
55	112863	T03148	Hs.4610	EST	0.092
	129290	AA521407	Hs.110095	ESTs	0.092
	103384	X92762	Hs.79021	ESTs	0.092
				tafazzin (cardiomyopathy; dilated 3A (X-linked); endocardial fibroelastosis 2; Barth syndrome)	0.092
				ESTs	0.092
60	112508	R68213	Hs.28847	ESTs	0.092
	111863	R37495	Hs.23578	ESTs	0.092
	131184	AA452705	Hs.23954	ESTs; Weakly similar to KIAA0584 protein [H.sapiens]	0.092
	107420	W26567	Hs.4775	ESTs	0.092
	111768	R27606	Hs.24185	ESTs	0.092
65	112290	R53940	Hs.26016	ESTs	0.092
	130581	AA481982	Hs.16258	ESTs; Weakly similar to RAS-RELATED PROTEIN RAB-5A [H.sapiens]	0.092
	120744	AA302772	Hs.228649	EST	0.093
	112226	R50761	Hs.25738	ESTs	0.093
	116154	AA460951	Hs.57100	ESTs	0.093
	102640	U67674	Hs.194783	solute carrier family 10 (sodium/bile acid cotransporter family); member 2	0.093
	129797	X53595	Hs.1252	apolipoprotein H (beta-2-glycoprotein I)	0.093
	102705	U77180	Hs.50002	small inducible cytokine subfamily A (Cys-Cys); member 19	0.093
	132408	AA035547	Hs.47822	KIAA0380 gene product; RhoA-specific guanine nucleotide exchange factor	0.093
	108441	AA079079		zm97c9.s1 Stratagene colon HT29 (#937221) Homo sapiens cDNA clone	0.093

				IMAGE:545872 3' similar to contains element MER22 MER22 repetitive element ;, mRNA sequence	0.093
	108145	AA054133	Hs.63085	ESTs	0.093
5	106466	AA449990	Hs.76057	lysophospholipase II	0.093
	101697	M64358		Human rhom-3 gene, exon	0.093
	121294	AA401958	Hs.240170	ESTs; Moderately similar to alternatively spliced product using exon 13A [H.sapiens]	0.093
	117824	N49065	Hs.125201	ESTs; Weakly similar to B7 [M.musculus]	0.093
10	115771	AA422049	Hs.40780	ESTs	0.093
	102303	U33053	Hs.2499	protein kinase C-like 1	0.093
	131405	U79255	Hs.26468	amyloid beta (A4) precursor protein-binding; family A; member 2 (X11-like)	0.093
	112909	T10069	Hs.101094	ESTs	0.093
	124173	H41281	Hs.107619	ESTs	0.093
	112488	R66896	Hs.28788	ESTs	0.093
15	130554	X59303	Hs.159637	valyl-tRNA synthetase 2	0.093
	106413	AA447964	Hs.6311	ESTs	0.093
	111711	R22891	Hs.7093	ESTs	0.094
	117595	N34933	Hs.44664	EST	0.094
	113813	W45174	Hs.31382	ESTs	0.094
20	107769	AA018449	Hs.125220	Homo sapiens DNA from chromosome 19-cosmids R30102:R29350:R27740 containing MEF2B; genomic sequence	0.094
	114966	AA250743	Hs.92198	ESTs; Highly similar to calcium-regulated heat stable protein CRHSP-24 [H.sapiens]	0.094
	130297	H94949	Hs.171955	trophinin-assisting protein (tastin)	0.094
25	109589	F02429	Hs.6581	ESTs	0.094
	112592	R77631	Hs.29126	ESTs	0.094
	102314	U34038	Hs.154299	coagulation factor II (thrombin) receptor-like 1	0.094
	116128	AA459915	Hs.112193	mutS (E. coli) homolog 5	0.094
30	106809	AA479704	Hs.220324	Human DNA sequence from clone 283E3 on chromosome 1p36.21-36.33. Contains the alternatively spliced gene for Matrix Metalloproteinase in the Female Reproductive tract MIFR1; -2; MMP21/22A; -B and -C; a novel gene; the alternatively spliced CDC2L2 gene for	0.094
	130607	AA043894	Hs.16603	ESTs	0.094
35	120592	AA281929	Hs.143974	ESTs	0.094
	117230	N20535	Hs.43265	melastatin 1	0.094
	105948	AA404597	Hs.7133	ESTs	0.094
	101333	L47738	Hs.80313	p53 inducible protein	0.094
	101909	S69265		Homo sapiens mRNA for PLE21 protein; complete cds	0.094
40	106959	AA497031	Hs.8657	ESTs; Highly similar to CTG7a [H.sapiens]	0.094
	127034	AA352389		ESTs; Wkly smlr to glucose-6-phosphatase catalytic subunit [R.norvegicus]	0.095
	134430	H52105	Hs.8309	KIAA0747 protein	0.095
	120342	AA207105	Hs.45068	Homo sapiens mRNA; cDNA DKFZp434l143 (from clone DKFZp434l143)	0.095
	104450	L77564	Hs.103978	serine/threonine kinase 22B (spermiogenesis associated)	0.095
45	130902	AA424530	Hs.21061	ESTs	0.095
	102708	U77594	Hs.37682	retinoic acid receptor responder (tazarotene induced) 2	0.095
	107373	U85773	Hs.154695	phosphomannomutase 2	0.095
	123569	AA608952	Hs.195292	ESTs; Weakly similar to RNA helicase HDB/DICE1 [H.sapiens]	0.095
	102687	U73379	Hs.93002	ubiquitin carrier protein E2-C	0.095
	128888	AA034951	Hs.106893	ESTs	0.095
50	100283	D43642	Hs.2430	transcription factor-like 1	0.095
	102747	U79303	Hs.82482	protein predicted by clone 23882	0.095
	107798	AA019346	Hs.60918	EST	0.095
	123565	AA608907	Hs.112614	EST	0.095
55	116010	AA449450	Hs.56421	ESTs; Weakly similar to Similarity to H.influenza ribonuclease PH [C.elegans]	0.095
	117155	H97536	Hs.42391	EST	0.095
	133094	AA115572	Hs.64746	chloride intracellular channel 3	0.095
	113174	T54659	Hs.9779	ESTs	0.095
	102016	U03270	Hs.122511	centrin; EF-hand protein; 1	0.095
60	130126	AB002318	Hs.150443	KIAA0320 protein	0.095
	134813	X14767	Hs.89768	gamma-aminobutyric acid (GABA) A receptor; beta 1	0.095
	132055	N69440	Hs.38132	ESTs	0.095
	122229	AA436198	Hs.103902	ESTs	0.096
	127574	AA907314	Hs.188905	ESTs	0.096
	134432	AA053022	Hs.8312	ESTs	0.096
65	128052	AA878398	Hs.190491	ESTs	0.096
	101637	M58285	Hs.132834	hematopoietic protein 1	0.096
	103386	X92972	Hs.80324	protein phosphatase 6; catalytic subunit	0.096
	133079	AA477561	Hs.6449	ESTs	0.096
	120328	AA196979	Hs.104129	ESTs; Weakly similar to protease [H.sapiens]	0.096

5	107640	AA009615	Hs.257808	ESTs	0.096
	123389	AA521176	Hs.221231	ESTs	0.096
	103222	X74795	Hs.77171	minichromosome maintenance deficient (S. cerevisiae) 5 (cell division cycle 46)	0.096
	111704	R22450	Hs.23396	ESTs; Highly similar to ZINC FINGER PROTEIN 140 [H.sapiens]	0.096
	126856	AA306523		EST177475 Jurkat T-cells VI Homo sapiens cDNA 5' end, mRNA sequence.	0.733
10	127071	AA250806		ESTs	0.096
	114550	AA056755	Hs.151714	ESTs	0.096
	125955	AI356943	Hs.143761	ESTs	0.096
	134363	M37033	Hs.82212	CD53 antigen	0.096
	128550	W76492	Hs.170142	ESTs	0.096
15	122598	AA453465	Hs.99329	ESTs	0.096
	118898	N90703	Hs.4236	KIAA0478 gene product	0.096
	117661	N39092	Hs.44940	ESTs	0.096
	120996	AA398281	Hs.143684	ESTs	0.096
	123388	AA521172	Hs.134417	ESTs	0.096
20	106700	AA463929	Hs.28701	ESTs	0.096
	112962	T16814	Hs.6828	ESTs	0.096
	121262	AA401372	Hs.97723	ESTs	0.096
	134551	R44839	Hs.8526	l-beta-1;3-N-acetylglucosaminyltransferase	0.096
	112060	R43754	Hs.21164	ESTs	0.096
25	134678	AA039935	Hs.182595	dynein; axonemal; light polypeptide 4	0.096
	100855	HG4234-HT4504		Methylenetetrahydrofolate Reductase	0.097
	132414	N91193	Hs.48145	ESTs	0.097
	112900	T08758	Hs.3813	ESTs	0.097
	115989	AA447777	Hs.93135	ESTs	0.097
30	103561	Z21488	Hs.143434	contactin 1	0.097
	131087	AA009738	Hs.22824	ESTs; Weakly similar to p160 myb-binding protein [M.musculus]	0.097
	120293	AA190859	Hs.191428	ESTs	0.097
	111830	R36081	Hs.25085	EST	0.097
	113654	T95770	Hs.17666	ESTs	0.097
35	132675	AA179338	Hs.5476	serine proteinase inhibitor	0.097
	120182	Z40125	Hs.91968	ESTs	0.097
	132679	U16282	Hs.5881	ELL gene (11-19 lysine-rich leukemia gene)	0.097
	134211	AA056681	Hs.80021	ESTs; Weakly similar to 62D9.p [D.melanogaster]	0.097
	115448	AA284845	Hs.165051	ESTs	0.097
40	118118	N56901	Hs.47995	ESTs	0.097
	107598	AA004528	Hs.169444	ESTs	0.097
	128933	H01824	Hs.760	GATA-binding protein 2	0.097
	114892	AA235988	Hs.86024	ESTs	0.097
	101922	S75168	Hs.274	megakaryocyte-associated tyrosine kinase	0.097
45	105444	AA252374	Hs.19333	ESTs; Weakly similar to ATP(GTP)-binding protein [H.sapiens]	0.097
	128155	AA926843	Hs.143302	ESTs	0.097
	116276	AA485870	Hs.44914	ESTs	0.097
	111964	R41227	Hs.21860	ESTs	0.097
	135100	AA398926	Hs.251108	Homo sapiens mRNA; chromosome 1 specific transcript KIAA0493	0.097
50	124872	R69251	Hs.101506	EST	0.097
	103084	X59932	Hs.77793	c-src tyrosine kinase	0.097
	124138	H23199	Hs.107010	ESTs	0.098
	130048	R31745	Hs.211612	SEC24 (S. cerevisiae) related gene family; member A	0.098
	100208	D26129	Hs.78224	ribonuclease; RNase A family; 1 (pancreatic)	0.098
55	123537	AA608775	Hs.112589	ESTs	0.098
	118999	N95019	Hs.55092	ESTs	0.098
	119847	W80384	Hs.9853	ESTs	0.098
	112819	R98618	Hs.35984	ESTs	0.098
	131080	J05008	Hs.2271	endothelin 1	0.098
60	127353	AA190853	Hs.155360	ESTs	0.098
	132068	X66365	Hs.38481	cyclin-dependent kinase 6	0.098
	105744	AA293436	Hs.12909	ESTs	0.098
	133680	M92357	Hs.101382	tumor necrosis factor; alpha-induced protein 2	0.098
	122899	AA469960	Hs.178420	ESTs; Highly similar to WASP interacting protein [H.sapiens]	0.098
65	128700	U59286	Hs.103982	small inducible cytokine subfamily B (Cys-X-Cys); member 11	0.098
	104393	H46486	Hs.226499	nesca protein	0.098
	123320	AA496792	Hs.139572	EST	0.098
	129169	N31641	Hs.109058	ribosomal protein S6 kinase; 90kD; polypeptide 5	0.098
	135093	U51333	Hs.159237	hexokinase 3 (white cell)	0.098
	113269	T65159	Hs.85044	ESTs	0.098
	124283	H86783	Hs.194136	ESTs; Moderately similar to zinc finger protein RIN ZF [R.norvegicus]	0.098
	114376	GMCSF		Accession not listed in Genbank	0.099
	100881	HG4458-HT4727		Immunoglobulin Heavy Chain, Vdjc Regions (Gb:L23563)	0.099

5	116572	D45654	Hs.65582	DKFZP586C1324 protein	0.099
	123956	AA621747	Hs.112847	EST	0.099
	100818	HG4018-HT4288		Opioid-Binding Cell Adhesion Molecule	0.099
	132754	W47419	Hs.56007	Human DNA from chromosome 19-specific cosmid F25965; genomic sequence	0.099
	112741	R93080	Hs.35035	ESTs	0.099
10	112748	R93299	Hs.166492	ESTs	0.099
	130858	S57235	Hs.246381	CD68 antigen	0.099
	124870	R69233	Hs.101504	ESTs	0.099
	125304	Z39833	Hs.124940	GTP-binding protein	0.099
	121297	AA401995	Hs.97860	ESTs	0.099
15	128602	AA046103	Hs.102367	ESTs	0.099
	124062	H00440	Hs.144524	ESTs; Weakly similar to signal transducer and activator of transcription 2 [M.musculus]	0.099
	100547	HG2149-HT2219		Mucin (Gb:M57417)	0.099
	105652	AA282505	Hs.19015	ESTs	0.099
	133390	AA459945	Hs.72660	KIAA0585 protein	0.099
20	133503	M33195	Hs.743	Fc fragment of IgE; high affinity I; receptor for; gamma polypeptide	0.099
	109461	AA232667	Hs.58210	ESTs	0.099
	102068	U09117	Hs.80776	phospholipase C; delta 1	0.099
	113464	T86931	Hs.16295	ESTs	0.099
	104240	AB002368	Hs.70500	KIAA0370 protein	0.099
25	121113	AA399109	Hs.161813	ESTs	0.1
	122896	AA469952	Hs.97899	ESTs; Weakly similar to dal2; len:343; CAI: 0.17; ALC_YEAST P25335	0.1
	102405	U43148	Hs.159526	ALLANTOICASE [S.cerevisiae]	0.1
	103599	Z33905	Hs.81218	patched (Drosophila) homolog	0.1
	121079	AA398719	Hs.14169	receptor-associated protein of the synapse; 43kD	0.1
30	115820	AA427487	Hs.39619	ESTs; Weakly similar to CREB-binding protein [H.sapiens]	0.1
	125106	T95766	Hs.189760	ESTs; Weakly similar to RETICULOCALBIN 1 PRECURSOR [H.sapiens]	0.781
	131373	N68116	Hs.26146	ESTs	0.1
	120224	Z41239	Hs.106960	Down syndrome critical region gene 3	0.1
	133090	AA448228	Hs.6468	ESTs	0.1
35	132300	AA133244	Hs.44234	ESTs	0.1
	113129	T49384	Hs.8988	EST	0.1
	110638	H73197	Hs.17241	ESTs	0.1
	131364	R53255	Hs.26010	ESTs	0.1
	105370	AA236476	Hs.22791	ESTs; Weakly similar to transmembrane protein with EGF-like and two follistatin-like domains 1 [H.sapiens]	0.238
40					

TABLE 11A shows the accession numbers for those primekeys lacking unigeneID's for Table 11. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey:	Unique Eos probeset identifier number
CAT number:	Gene cluster number
Accession:	Genbank accession numbers

Pkey	CAT number	Accession
100610	19864_1	AW161357 AI879062 AI928938 AW161097 AW161167 BE314465 AA351715 F07096 AA179034 F08510 F00653 AI936671 AA476718 AW772454 AI807703 R44253 AA976667 AI985186 AI650254 H38942 R84829 AA018724 AA001000 H85934 AA019126 H85609 AA017000 AA339355 AW950556 D51397 AA213981 BE548002 AI056359 AA001560 AW952113 AA317769 AI857477 AI857475 AW249771 AW162661 H38943 AA018628 R85885 AI984613 AI934765 AI796172 AW157488 AI929191 R85523 D51221 D53851 H85610 AI749674 F21582 AA323145 AA019127 AA687444 T06745 AI699293 H29532 AA214029 AA223656 NM_016834 X14474 R19697 H09695 R17455 R13812 R19056 AI681231 AI590200 R37671 AA861828 AI990023 AI935669 AW005821 AA324581 H17335 R37659 R42802 R46242 R60936 R59731 H28993 AA479907 R44570 AI890696 AA308884 AA507078 R41274 AI365507 T16348 AI560453 F03259 F04722 T16312 AA016081 AW073061 BE314824 W28930 R44098 R51045
100674	21517_2	AW403342 AW248986 BE561709 AA357312 BE311834 BE389496 BE294887 AW732696 BE047868 AI702383 BE019155 AI702367 BE408966 BE280458 BE313759 BE513492 BE535404 BE280258 AC005263 NM_007165 L21990 AW732711 AI564920 AW249094 BE265365 AW607186 AW607346 BE005217 H27211 U46230 BE260066 BE207043 BE546782 AW248659
108559	41469_9	AA085228 AA085161
100721	19818_1	L40904 NM_005037 X90563 AB005526 H21598 AA088517
100748	41861_1	X06096 X05826
100750	15759_1	BE157260 BE157265 R48118 H43827 Z17877 AW379070 AW291778 M20605 J03253 M14206 V00568 AI860465 AW296022 M13930 AL047400 J00120 BE018476 AW675223 T26980 F06694 R22709 R24720 H22753 AI903100 AI903094 AW937823 X00364 D10493 K01904 K01906 K00535 L00058 AA410662 AW384760 AA304930 AI680985 X00198 H58025 AW998901 AV653447 N31654 AW610357 AW610369 AW862480 BE223010 AW384172 AW384219 AW384171 AW384218 AA298522 BE140421 AW945162 AW751711 AA514409 AW747912 AI214214 W87741 AA972406 AA554513 BE302087 AI249030 AA477850 AV653129 AI281360 AI274110 W87861 AA641366 X66258 AI051600 AA877139 AA527483 AA857219 AI250782 AA625531 AA807892 AI278811 AI224033 H24033 AA593396 AW129709 R45453 N22772 AA235530 T29737 AI016409 AI688907 AA568370 AA722760 AI539329 AA550843 AW674698 AI538452 AI538453 AI337957 AA477744 AA464600 AI140319 AW949294 AI339781 AI828736 AA923634 AA344094 AI278350 AA975567 AA908416 AA857170 AW023520 R43413 R48004 F02958 AI989439 R11207 AA737307 D10493 AW950652 AI093842 AI474024 AW703369 R11264 M13930 M13930 M13930 M13930 J00120 M13930 M13930 X00364 J00120 R19507 AA639812
100751	24700_1	N32759 N29730 N30831 N32604 N31955 AI206390 H87574 R23494 AI186215 N30036 AI741512 J00117 NM_000737 AI453626 AA330974 AI188729 AI188604 AI188964 N30276 AI188947 AI188830 AI188303 AI200457 AI219166 AI192459 AI183280 AI189275 AI188639 AI186353 AI189616 AI184224 AI130720 AI188454 AI188391 AI148857 AI192447 AI209155 AI190013 AI206355 AI188721 AI189429 AI189364 AI186330 AI431595 AI189595 AI188781 AI148647 AI200022 AI221552 AI220923 AI188728 AA233034 AI189807 AI189641 AI219044 AI148774 AI200658 W71989 AI207360 AI188824 AI200559 AI200270 AA644163 AI199943 AI151301 AI189555 AI262724 AI148590 AI148695 AI126906 AI149163 K03183 K03189 AI189842 AI221014 N30608 AI186465 AI220865 AI188498 AI138226 AI189968 AI221019 AI138197 AI149426 AI148904 AI186218 AI188348 AI160579 AI198460 AI149039 AI160936 AI219055 AI184784 AI221580 AI161082 AI160814 AI123896 AI417614 AI126101 AI188872 AI149571 AI168533 AI149072 AI149467 AI131286 N30684 AI160705 AI160692 AI149559 AI273580 AI189442 AI138448 AI149591 N27302 AA400910 AI138431 AI138435 AI128407 N30216 AI128296 AI219589 AI188492 AI149447 AI168482 H95374 AI219009 N31616 AI276216 N32233 AI291937 N30741 AI188689 N27111 R23214 AI221605 AI184348 AI200375 H94451 N26397 AI871881 AA232905 N30833 AI220780 H94446 N30822 H87464 R68815 N30290 AI128424 H12587 T47334 H87631 H87156 AI219133 AI868741 AA330859 H86993 AA330413 H93656 N30817 T90191 H93668 AI200054 H95207 T47316 H95381 T49170 R00880 T49171 N27381 H94107 R63352 T85053 AW451899 H95142 N30313 H94015 H86987 T28278 N29701 C18834 AA331267 AA330939 AI654493 N27073 N29831 R68113 N30758 R26086 N32108 H95135 AA330414 AA330978 AI219422 AI189453 AI199951 X00264 NM_000894 AA371909 AA063496 T29543 AA371971 AA372026 AA371978 AA371346 AI051683 AI186418 AI220659 AI189068 AI219266 AI186552 AI188715 AI149156
100760	1334_7	AW794626 M27126 M27014
100775	18179_3	J05581 M61170 T27692 M34088 M34089 AW860335 AW579047 AW610437 AW610386 AW610422 AW610473 AW579078 AW604897 AW860163 AW579067 AW862410 AI816584 AW177757 AW602769 AI909790 AW860331 AI909787 AI909811

5

10

15

20

25

30

35

40

45

50

55

60

65

AI909813 AW845083 AI905920 AW387919 BE140766 AI909279 AW369405 AA429321 AA429320 AA367451 AA847972
 AW001137 AI567905 T84561 AI631295 AA151351 H02932 AI884519 AA367457 AW369421 AI678846 AW391803 AI610869
 AW192838 AI922289 AI952140 AI910233 AI479474 AW001395 AA488073 AI985760 AW130017 AI858369 AA627845
 AW081805 AA158865 AI624443 AA344985 AA569793 R72486 AI589329 AI903204 AI269893 AA641284 AI279932 AA149270
 AI697120 AA729146 AI589353 AA480067 AI923310 AA530908 AI275395 AA425062 AA580280 AA889527 AA158866
 AW131341 AA573028 AA877326 T29335 AW951288 H04235 AA099243 AA994659 AI659618 AA887919 AI299297
 AW001116 AW263844 AI270578 AA970828 AW572126 AA775299 AW369449 AW369398 AW369452 AI933677 AI870710
 AI092911 AI582464 AI497674 AA937026 AA885865 L38597 AA908325 AW369432 AW026623 AA627778 AI264942
 AA932409 AI187328 AI672970 AI886098 AW440471 AW138860 AI866858 AI802528 AI926172 AW243914 AI933690
 AA996114 AA536189 AW009937 AI918060 AI270379 AI973169 AW175638 AW369413
 NM_006227 L26232 R50649 AU077024 AL008726 AA411079 R35151 BE278153 BE278139 AI459777 R88036 Z43210
 F07326 AF052157 R17844 BE615476 T82160 R71985 H21963 AA299158 AW368246 R48123 R50628 R70441 H27245
 H72015 R72345 R39392 AI909738 BE612778 BE613234 D52116 D52136 D52132 D52067 D51922 D51995 D51905 N34249
 N25459 AA464436 AA297350 AA297466 R81736 H02737 AW582505 R27523 AI834241 AW130867 W72668 W76426
 AA358363 R50262 AW473860 H52335 H43953 H21964 T39505 AI887517 AW156925 AW839850 H02628 AW007705
 AI561008 F22392 R71279 AA995433 R50725 W24462 R71931 AA464437 AW591731 R25667 R25695 R50810 AI560805
 AI089266 H68386 H41353 H28590 AW001860 AI141623 AA250773 AI284778 AW511412 AW083975 AA130377 AW026047
 R50551 R81494 AI357668 AI078272 F32666 F36981 AW304865 H43906 AA931068 R48010 AI540217 AI017339 AI291812
 AI741954 AA458490 AI088378 AA298764 H61168 AA358362 AA298725 AA298515 AA464148 AA443538 R43046 AA084314
 T40641 T47608 T48940 AI082477 AW470145 N92284 AI758958 AA298512 AA284586 AI597777 AA480277 AI932559
 AI869081 AA476615 AA503651 AI656024 AW168522 AI682051 AI689106 AI274592 AI520917 BE258916 BE615861
 BE280282 R53386 BE278255 BE278398 T47607 AA477662 H68385
 100817 19648_1 L34355 L46810 NM_000023 U08895 AA424260 AI097272 AA424162 N79764 F19290 F25278 AI479385
 AA460662 AA432059 AW016935 F25770 F32549 F36677 F33016 F35992 F36010 AW172497 AA835076 F28727 AA211643
 AA453282
 100818 19604_3 U79251 AA843851 R38201 R66461 R44908 AA683289 H17477 R37364 R52832 AW298336 AA351391 NM_002545 L34774
 AA296886 AW967001 T28889 R13451 T77331 AL119196 AL118830 H08459 AW892812 AW905838 H17585 R52878
 100881 458_127 BE561958 BE561728 BE397612 BE514391 BE269037 BE514207 BE562381 BE514256 BE514403 BE514250 BE397832
 BE269598 BE559865 BE396881 BE560031 BE514199 BE560037 BE560454
 100885 12707_3 X07881 NM_006249 X07637 AA376715 AA376677 X07715 X07704 S80916
 100898 8542_1 BE387614 R51501 AA199714 AW674779 F08178 BE269071 AA376313 H08264 AA380420 H18785 AL042151 BE277758
 BE267438 NM_005850 L35013 BE540833 BE390902 BE391494 BE277459 BE385592 BE390612 BE384263 BE387779
 BE388647 BE537373 BE547158 AW409585 AW374033 AW602185 AA355725 AW577548 AW935015 AW935160 W40232
 AW938647 AW374332 AA434040 BE293488 AL138361 BE560260 AI745075 AA317980 AW949382 AI834311 AI653582
 AI831042 AI361878 AA618606 AA729052 AI424969 AA199715 AW769374 AI828422 AW044307 AI862816 AI203583
 AW084461 AW514655 AA831883 AA290672 AA831286 AA578510 AW089965 AW150746 AA292743 H22232 AI469275
 AW439312 AA292744 AW471443 AI473989 AA593336 AA464070 AI678937 AW069451 AA970763 AA610480 AA593328
 AA464009 AA768895 AI298928 AA436600 AA464718 AA699361 D61482 D55935 AI369591 AA470695 AI809135 AA640627
 AI568446 R51502 W45467 AI655316 AA463934 AW168609 AW518663 BE045525 Z41251 AI868091 AA908160 AI026697
 AI886259 AI612932 AA215437 AI956014 BE541087 BE255652 BE265878 BE394102 W27502
 102459 3556_1 U48936 L36592 X87160 NM_001039 AL036606 AL036420 U35630 AW298574
 126126 1630017_1 W80551 M85370
 102620 16821_37 AA976427 U66052
 102673 24986_6 AI457548 U72509
 102675 5145_4 U72512 T98357 R31335 F18090
 102753 2226_1 L32961 NM_000663 U80226 S75578 AA425061 AA429317 AI815143 AA910669 AI286022 AI286019
 102799 34624_4 U88896 U88898 AA916056 T03285 AI341594 AI359534 AI634031 U88897
 127034 51148_2 BE397750 AA232171 BE562900 BE384894 BE242228 BE206819 BE261742 AA296468 AW959763 BE276164 BE264109
 BE392626 BE256735 AA301453 N55872 H01676 AA292746 AA427485 AA496400 AA352389
 103522 21640_1 Y10518 Y10514 Z83935 Y10508 AK000055 Y10519 AI142012 AI681175 BE222219 AA890586 BE504347 BE328064 N63044
 N51226 AI151248 AI521996 AI924777 AW375954 AI860275 W00549 AI742673 AW612288 AI763062 AA632510 AI087347
 AI088070 AI214349 AA890297 AI494156 AI698598 AA631658 AA504593 AA860733 AI266761 AW663214 AW771231
 AA639610 AI769806 AI769746 AW014326 AI288611
 127071 188097_1 AA250806 AA459220
 126456 291965_1 AA429212 W00881
 119388 1762256_1 T88798 R92430
 126856 20669_1 AI084125 AI083773 AI479687 AI939609 AI968662 AF129507 NM_013282 AW971840 AW298508 AA744240 AA811217
 AA827671 AA811055 AA806567 AA488977 AA908902 AI637637 AA927056 AI870139 AW340492 AA488755 AA129794
 AA306523 AA354253 BE256277 AC053467 AW962084
 103996 224545_1 AA321355 AW964592 R23284 H73883 R23382 N47914 C01377 H04668 AW606248 R34447 AA847136 AI684489 AI523112
 AW044269 AI379138 N29366 AA761543 N79248 AA960845 AA768316 AI147926 AI718599 AI880620 R67467 AI216016
 AI738663 H04648
 113213 23798_1 NM_001395 Y08302 AI434619 AI470328 AI261807 AW024965 AI806537 AI830549 AI640337 AI219065 AW271700
 AW028488 AI133339 AI859205 R51175 U87167 BE379324 BE392008 AA340819 AA343110 T57275 D59164 AW299312
 AI434422 AI936390 AW024975 R40262
 134947 844579_1 AW269126 R09430 T56590 AI367247 AI253132 BE464248 T58658 AW207785 T58607
 129311 16078_1 R51194 AI732276 R53587 AI820697
 AK000526 BE550084 W30689 AW271859 AA411456 AI341551 AA242990 AA243027 H87046 D20360 AI184053 AA146956
 AI721023 AI718944 AA146955 F18215 AA903890 AI700355 AI075430 AA411584 AA878210 AI476760 AW945637 AA630596

		AA431522 AA301989 AI909058 D12149 N41960 BE222214 AA609922 AA828176 AA393359 AA398693 AW024956 BE467805 AW298623 AW264085 AI024454 AI024719 AI431927 T55087 AI611014 T54920 AA131253 AI436344 AA017176 AI359979 AA047836 AA017063 AA016303 AA001545
5	114427 9724_2 114569 110077_1 100106 15621_5 100515 342_1	AA063315 AA063316 AF015910 AA305746 D90187 T63943 AW951154 T29182 AI734941 D13264 AI299239 Z18812 AW299859 W24476 AA933064 AA489759
10	100531 46038_1 100545 22955_11 100574 17320_2	AW888554 AW607282 AA319986 M28590 M55405 AW752552
15		AA326895 M10036 NM_000365 N84665 H69414 N84657 AA380453 AA329743 AA357367 AA188770 AA376532 AA353653 AA158953 AA083176 BE537313 AA181433 D53373 R57376 AA206698 R14807 H18899 H11191 H93892 R25593 T61134 N93285 AA083081 AA831789 H13137 AA497014 AA079330 AA182861 H13138 W47161 R62913 AA687089 AA211112 AA429237 AL035923 AA100070 AW392898 AI566433 AA866006 AA214002 AW392865 N79454 AA197181 AI680371 AA176501 AA737967 AI089225 F34874 AW571437 AI620620 AA573489 AA423816 AA164917 AA458455 T47072 AI569087 AI261656 AA730919 AI633441 AW195182 AI351622 AW243465 AI872649 AI359227 AA987941 AI693770 T47073 AW779948 AW510580 AI635626 AW627601 AA864326 AA953578 AI341418 BE222853 AI241963 AI094663 AA928380 AA493373 AW043762 AI377783 AW958987 BE619760 AA385240 BE277975 BE280095 AW631443 AA581048 BE618715 BE299610 C14874 BE559858 BE378455 BE618290 BE544585 AI525575 BE548897 BE267110 AA804738 BE269821 AA918133 BE277647 AA599947 BE280735 BE390239 N74150 T12504 AI208197 AW955527 AA113897 N40081 H73835 H70393
20		AI434041 W22950 AI192661 BE264461 W26486 AA626424 AA196694 T69209 AA857976 AI540287 AA410599 AA864287 AW950564 AA013320 T49283 AI541438 AW804703 AA335534 AA335659 BE562269 BE618802 BE277850 BE546413 BE280994 AA204813 BE561694 BE543524 BE253647 AW001452 W19116 BE542508 AA205894 BE254875 BE270033 AI525906 BE251792 AA975700 BE272138 AW607671 N87686 M10036 BE515060 BE298607 AI745178 U47924 H03193
25	100627 tigr_HT2798 100756 tigr_HT3768 100768 tigr_HT3846 100813 tigr_HT4265 100836 tigr_HT4383 100855 tigr_HT4504	Z25424 M88357 L29141 M69180 M81105 L33999 U04688 U09806
30	102104 entrez_U12139 125091 genbank_T91518 100929 tigr_HT688 125147 _entrez_W38150 102354 entrez_U38268 102491 entrez_U51010 102636 entrez_U67092 118769 genbank_N74496 101046 entrez_K01160 101057 entrez_K03430	U12139 T91518 X65561 W38150 U38268 U51010 U67092 N74496 K01160 K03430
40	108334 genbank_AA070473 108417 483241_1 108441 genbank_AA079079 108786 genbank_AA128999 101655 entrez_M60299 101697 entrez_M64358 117437 genbank_N27645 101798 entrez_M85220 101909 entrez_S69265 103508 entrez_Y10141 103575 entrez_Z26256 119332 genbank_T54095 112161 genbank_R48295 119564 NOT_FOUND_entrez_W38206 114376 NOT_FOUND_entrez_GMCSF	AA070473 AA070853 AA075749 AA075716 AA079079 AA128999 M60299 M64358 N27645 M85220 S69265 Y10141 Z26256 T54095 R48295 W38206 GMCSF
55	100478 tigr_HT1067 100547 tigr_HT2219 100564 tigr_HT2324	M22406 M57417 Z11585

TABLE 12: shows genes, including expression sequence tags, that are down-regulated in prostate tumor tissue compared to normal prostate tissue as analyzed using Affymetrix/Eos Hu01 GeneChip array. Shown are the ratios of "average" normal prostate to "average" prostate cancer tissues.

10	Pkey:	Unique Eos probeset identifier number			
	ExAccn:	Exemplar Accession number, Genbank accession number			
	UnigeneID:	Unigene number			
	Unigene Title:	Unigene gene title			
	R1:	Background subtracted normal prostate : prostate tumor tissue			
15	Pkey	ExAccn	UnigeneID	Unigene Title	R1
	100522	HG1763-HT1780		Prolactin-Induced Protein	17.4
	130803	M81650	Hs.1968	semenogelin I	16.785
	118068	N53943	Hs.13743	ESTs	13.225
20	114251	Z39898	Hs.21948	ESTs	12.7
	112134	R46025	Hs.7413	ESTs	8.735
	101436	M20642	Hs.158295	Human alkali myosin light chain 3 mRNA; complete cds	8.175
	104028	AA361094	Hs.221128	ESTs	8.15
	108944	AA149204	Hs.175783	ESTs; Highly similar to growth arrest inducible gene product [H.sapiens]	7.535
25	103838	AA174173	Hs.12622	ESTs	7.212
	120469	AA251741	Hs.25882	DKFZP586M1824 protein	7.175
	110279	H29231	Hs.27384	ESTs	6.701
	127472	AA761378	Hs.192013	ESTs	6.642
	133301	N35229	Hs.7037	pallid (mouse) homolog; pallidin	6.411
30	102457	U48807	Hs.2359	dual specificity phosphatase 4	6.395
	114011	W90385	Hs.15082	ESTs	6.15
	101249	L33881	Hs.1904	protein kinase C; iota	6
	123265	AA491209	Hs.105265	ESTs; Weakly similar to reverse transcriptase [M.musculus]	6
	119322	T49655	Hs.241569	ESTs; Modly smlr to !! ALU SUBFAMILY SQ WARNING ENTRY !! [H.sapiens]	5.95
35	101673	M61906	Hs.6241	phosphoinositide-3-kinase; regulatory subunit; polypeptide 1 (p85 alpha)	5.925
	115586	AA399218	Hs.92423	ESTs	5.7
	120590	AA281780	Hs.111441	ESTs; Weakly similar to similar to Kruppel-like zinc finger protein [C.elegans]	5.7
	109748	F10192	Hs.248323	Tubulin; alpha; brain-specific	5.625
	134727	X80507	Hs.8939	yes-associated protein 65 kDa	5.5
40	129171	AA234048	Hs.7753	calumenin	5.486
	120390	AA233122	Hs.111460	ESTs; Highly similar to multifunctional calcium/calmodulin-dependent protein kinase II delta2 isoform [H.sapiens]	5.4
	131699	R68657	Hs.90421	ESTs; Modly smlr to !! ALU SUBFAMILY SX WARNING ENTRY !! [H.sapiens]	5.279
	104490	N71503	Hs.43087	ESTs; Weakly similar to dysferlin [H.sapiens]	5.266
45	102124	U14528	Hs.29981	solute carrier family 26 (sulfate transporter); member 2	5.151
	109280	AA196635	Hs.86081	ESTs	5.134
	109707	F09739	Hs.185701	Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 21920	5.075
	108087	AA045709	Hs.40545	ESTs	5.075
50	135006	M21665	Hs.929	myosin; heavy polypeptide 7; cardiac muscle; beta	5.055
	119182	R80664	Hs.77067	ESTs	5.033
	129806	R62444	Hs.173373	KIAA0931 protein	4.675
	101435	M20543	Hs.1288	actin; alpha 1; skeletal muscle	4.626
	125954	R93943		yt72c12.r1 Soares retina N2b4HR Homo sapiens cDNA clone IMAGE:275735 5'	4.6
	113989	W87544	Hs.221184	ESTs	4.559
55	104432	J03460	Hs.99949	prolactin-induced protein	4.451
	112326	R56068	Hs.4268	ESTs	4.45
	119063	R16833	Hs.53106	ESTs; Weakly similar to !! ALU SUBFAMILY J WARNING ENTRY !! [H.sapiens]	4.45
	130376	R40873	Hs.155174	KIAA0432 gene product	4.301
	122484	AA448286	Hs.98074	ESTs; Highly similar to atrophin-1 interacting protein 4 [H.sapiens]	4.2
60	104142	AA447006		ESTs; Moderately similar to !! ALU SUBFAMILY SQ WARNING	4.175
	129413	N32787	Hs.11123	ESTs; Moderately similar to hypothetical protein 2 [H.sapiens]	4.1
	103678	Z84483		Human DNA sequence from PAC 46H23, BRCA2 gene region chromosome 13q12-13.05	
	114266	Z40186	Hs.26409	ESTs	4.05
	115206	AA262491	Hs.186572	ESTs	4.048
65	123723	AA609749	Hs.112759	ESTs; Highly similar to unknown protein [R.norvegicus]	4.041
	129130	H97993	Hs.172788	ESTs; Weakly similar to KIAA0512 protein [H.sapiens]	4.028

	120217	Z41078	Hs.66035	ESTs	4.028
	108536	AA084524		zn19d8.s1 Stratagene neuroepithelium NT2RAMI 937234 Homo sapiens cDNA	4.023
	134460	AA400030	Hs.8360	ESTs; Weakly similar to !! ALU CLASS B WARNING ENTRY !! [H.sapiens]	3.925
5	120418	AA236010	Hs.26613	Homo sapiens mRNA; cDNA DKFZp586F1323 (from clone DKFZp586F1323)	3.91
	132783	N74897	Hs.5683	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 15	3.889
	125052	T80174	Hs.222779	ESTs; Moderately similar to similar to NEDD-4 [H.sapiens]	3.85
	108600	AA099585	Hs.41175	ESTs	3.833
	103099	X61100	Hs.8248	NADH dehydrogenase (ubiquinone) Fe-S protein 1 (75kD) (NADH-coenzyme	3.818
10	134948	H06773	Hs.93850	protein kinase; AMP-activated; gamma 2 non-catalytic subunit	3.792
	120511	AA258144	Hs.221576	ESTs	3.779
	111861	R37460	Hs.25231	ESTs	3.768
	113966	W86600	Hs.9842	ESTs	3.75
	131649	AA481254	Hs.30120	ESTs	3.708
	129775	R94659	Hs.12420	ESTs	3.707
15	110191	H20568	Hs.27182	phospholipase A2-activating protein	3.7
	112678	R87160	Hs.33665	ESTs	3.7
	127115	AA375791	Hs.131894	ESTs	3.674
	132892	W92797	Hs.59378	DKFZP434G162 protein	3.653
	115023	AA252079	Hs.63931	dachshund (Drosophila) homolog	3.625
20	114932	AA242751	Hs.16218	KIAA0903 protein	3.62
	106865	AA487228	Hs.19479	ESTs	3.614
	134480	AA024664	Hs.83916	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex; 5 (13kD; B13)	3.613
	124780	R42493	Hs.220839	ESTs	3.6
	130631	AA025399	Hs.169737	ESTs	3.592
25	134154	AA211320	Hs.79404	neuron-specific protein	3.568
	104160	AA455706	Hs.99722	ESTs; Weakly similar to 78 KD GLUCOSE REGULATED PROTEIN PRECURSOR	3.559
	105524	AA258158	Hs.22153	ESTs; Weakly similar to KIAA0352 [H.sapiens]	3.542
	110168	H19673	Hs.176586	ESTs	3.525
30	109480	AA233299	Hs.72158	ESTs	3.522
	109585	F02367	Hs.27252	ESTs	3.5
	115134	AA257107	Hs.194331	ESTs	3.5
	116083	AA455653	Hs.44581	ESTs; Weakly similar to HEAT SHOCK 70 KD PROTEIN 6 [H.sapiens]	3.459
	120524	AA261852	Hs.192905	ESTs	3.45
35	116932	H74330	Hs.150000	ESTs	3.425
	130746	AA256976	Hs.18800	ESTs; Weakly similar to KIAA0579 protein [H.sapiens]	3.42
	107513	X05451	Hs.158295	Human alkali myosin light chain 3 mRNA; complete cds	3.417
	118641	N70298	Hs.49829	ESTs	3.407
	126584	AI028384	Hs.127331	ESTs	3.399
40	105134	AA159953	Hs.22895	ESTs; Weakly similar to arylsulfatase B precursor [H.sapiens]	3.325
	123502	AA600116	Hs.112526	ESTs	3.318
	132389	N50866	Hs.47135	ESTs	3.317
	105691	AA287097	Hs.75356	transcription factor 4	3.315
	131505	H85897	Hs.27755	ESTs	3.309
45	120775	AA342104	Hs.96777	EST	3.3
	105579	AA278824	Hs.19218	ESTs	3.295
	128190	AA946876	Hs.148376	ESTs	3.292
	100819	HG4020-HT4290		Transglutaminase	3.288
	130217	D29956	Hs.152818	ubiquitin specific protease 8	3.273
50	130068	AA608903	Hs.106220	KIAA0336 gene product	3.269
	134719	L07515	Hs.89232	chromobox homolog 5 (Drosophila HP1 alpha)	3.266
	110277	H29209	Hs.151231	ESTs; Highly similar to FYVE finger-containing phosphoinositide kinase [M.musculus]	3.26
	127354	AA418880	Hs.185797	ESTs	3.212
	129173	R60523	Hs.109087	ESTs	3.197
55	127464	AA970504	Hs.146103	ESTs	3.179
	124923	R94500	Hs.108046	ESTs	3.175
	122465	AA448164	Hs.99153	ESTs; Highly similar to CGI-73 protein [H.sapiens]	3.151
	122027	AA431302	Hs.98721	EST; Weakly similar to N-copine [H.sapiens]	3.151
	103329	X85134	Hs.72984	retinoblastoma-binding protein 5	3.15
60	129937	M95767	Hs.135578	chitinase; di-N-acetyl-	3.15
	134197	AA057341	Hs.87889	helicase-moi	3.15
	107764	AA018219	Hs.226923	ESTs	3.125
	121775	AA421773	Hs.161008	ESTs	3.125
	114768	AA149007	Hs.182339	Ets homologous factor	3.12
65	132381	N48818	Hs.46884	ESTs	3.11
	123105	AA485973	Hs.143947	ESTs	3.104
	121176	AA400080	Hs.97774	ESTs	3.1
	125053	T80620	Hs.186473	ESTs	3.075
	105909	AA401739	Hs.5111	ESTs	3.066

	119767	W72562	Hs.58119	ESTs	3.057
	115776	AA424038	Hs.58197	ESTs	3.056
	111713	R22988	Hs.220950	ESTs	3.05
5	115301	AA280047	Hs.43948	ESTs	3.05
	118448	N66412	Hs.49189	ESTs	3
	106586	AA456598	Hs.256269	ESTs	2.995
	110415	H48239	Hs.29739	ESTs; Weakly similar to RAS-RELATED PROTEIN RAB-3A [H.sapiens]	2.979
	105173	AA182030	Hs.8364	ESTs	2.978
10	101102	L07594	Hs.79059	transforming growth factor; beta receptor III (betaglycan; 300kD)	2.976
	110543	H58383	Hs.258544	ESTs	2.976
	125593	R24464	Hs.202949	KIAA1102 protein	2.964
	100824	HG4058-HT4328		Oncogene Aml1-Evi-1, Fusion Activated	2.957
	106822	AA481068	Hs.31835	ESTs	2.95
15	131963	D11930	Hs.3592	ESTs	2.95
	111221	N68869	Hs.15119	ESTs	2.936
	113620	T93795	Hs.17252	EST	2.917
	105220	AA210695	Hs.17212	ESTs	2.917
	123234	AA490227	Hs.105252	ESTs	2.904
20	125250	W87465	Hs.222926	ESTs; Weakly similar to D2092.2 [C.elegans]	2.9
	116196	AA465160	Hs.63386	ESTs	2.9
	122100	AA432243	Hs.41086	ESTs; Weakly similar to OXYSTEROL-BINDING PROTEIN [H.sapiens]	2.896
	111712	R22905	Hs.113716	ESTs	2.895
	126589	W78107	Hs.187698	ESTs; Weakly similar to Yer140wp [S.cerevisiae]	2.895
25	111132	N64378	Hs.13149	ESTs; Highly similar to unknown function [H.sapiens]	2.894
	115307	AA280300	Hs.191346	ESTs	2.886
	108989	AA152263	Hs.18827	KIAA0849 protein	2.883
	129486	H03686	Hs.220689	Ras-GTPase-activating protein SH3-domain-binding protein	2.879
	119805	W73788	Hs.43213	ESTs	2.875
30	125721	R59881	Hs.7503	ESTs	2.871
	103704	AA028171	Hs.153688	ESTs	2.868
	128420	AI088155	Hs.14146	ESTs; Weakly similar to unknown [H.sapiens]	2.866
	120571	AA280738	Hs.128679	ESTs	2.863
	123059	AA482019	Hs.238202	EST	2.86
35	129462	D84239	Hs.111732	IgG Fc binding protein	2.856
	125166	W45491	Hs.172609	nucleobindin 1	2.854
	125992	W01626		za36e07.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone	2.852
	109431	AA227972	Hs.43635	ESTs	2.85
	105077	AA142919	Hs.5558	ESTs	2.847
40	131388	R34531	Hs.92200	KIAA0480 gene product	2.846
	121080	AA398720	Hs.177953	ESTs	2.838
	112575	R73816	Hs.17385	ESTs	2.836
	130244	R26206	Hs.153293	KIAA0701 protein	2.825
	134698	AA427783	Hs.77910	3-hydroxy-3-methylglutaryl-Coenzyme A synthase 1 (soluble)	2.816
45	116355	AA504356	Hs.88650	ESTs	2.813
	115316	AA280627	Hs.57846	ESTs	2.806
	129677	U48736	Hs.198891	serine/threonine-protein kinase PRP4 homolog	2.8
	130971	H20332	Hs.28707	signal sequence receptor; gamma (translocon-associated protein gamma)	2.799
	115054	AA252863	Hs.87729	ESTs	2.795
	130285	AA063546	Hs.202968	ESTs	2.792
50	124308	H93575	Hs.227146	Homo sapiens mRNA; cDNA DKFZp564J142 (from clone DKFZp564J142)	2.783
	125502	AA732329	Hs.191959	ESTs	2.778
	114800	AA159825	Hs.131887	ESTs; Weakly similar to ORF YNL227c [S.cerevisiae]	2.768
	128625	AA242816	Hs.102652	ESTs; Weakly similar to KIAA0437 [H.sapiens]	2.766
	130159	H51098	Hs.151310	PDZ domain protein (Drosophila inaD-like)	2.75
55	107127	AA620504	Hs.22119	ESTs	2.742
	113547	T90746	Hs.15233	ESTs	2.734
	104639	AA004622	Hs.18214	ESTs	2.727
	127609	AA622559	Hs.150318	ESTs	2.726
60	106922	AA490964	Hs.10056	ESTs	2.725
	124825	R52088		yg85c3.s1 Soares infant brain 1NIB Homo sapiens cDNA clone	2.725
	124333	H98683	Hs.154054	ESTs	2.708
	117634	N36421	Hs.107854	ESTs; Weakly similar to SODIUM- AND CHLORIDE-DEPENDENT GLYCINE TRANSP	2.706
65	101609	M54927	Hs.1787	proteolipid protein 1 (Pelizaeus-Merzbacher disease; spastic paraplegia 2; uncomplicated)	2.704
	117142	H96908	Hs.42251	ESTs	2.7
	112602	R79147	Hs.203365	ESTs	2.695
	106828	AA481505	Hs.13797	ESTs	2.68
	124377	N25996	Hs.179833	ESTs	2.675

	101026	J04970	carboxypeptidase M	2.675
	124560	N66393	Hs.102754 ESTs	2.675
	124066	H02494	Hs.101615 ESTs	2.671
5	130281	R12777	Hs.15395 ESTs; Weakly similar to ARGINYL-TRNA SYNTHETASE [H.sapiens]	2.66
	110949	N49602	Hs.13308 ESTs	2.65
	111031	N54839	Hs.221085 ESTs; Highly similar to mediator [H.sapiens]	2.633
	121770	AA421714	Hs.11469 KIAA0896 protein	2.63
	134132	U32519	Hs.220689 Ras-GTPase-activating protein SH3-domain-binding protein	2.626
10	112424	R62452	Hs.191265 ESTs	2.625
	122544	AA451679	Hs.194410 ESTs	2.625
	134425	X90568	Hs.172004 titin	2.624
	111114	N63391	Hs.9238 ESTs	2.619
	116119	AA459242	Hs.44445 ESTs; Weakly similar to Kelch motif containing protein [H.sapiens]	2.615
	112079	R44164	Hs.23014 ESTs	2.6
15	123033	AA481271	Hs.193945 ESTs	2.591
	124196	H52617	Hs.144167 ESTs	2.586
	125873	H14437	yl25a04.r1 Soares breast 3NbHBst Homo sapiens cDNA clone	2.58
	117684	N40184	Hs.45050 ESTs	2.575
	134938	D30037	Hs.168326 phosphatidylinositol transfer protein; beta	2.575
20	131822	AA215647	Hs.200332 ESTs	2.568
	135185	U71203	Hs.96038 Ric (Drosophila)-like; expressed in many tissues	2.564
	117690	N40467	Hs.93834 ESTs	2.557
	118807	N78582	Hs.50732 protein kinase; AMP-activated; beta 2 non-catalytic subunit	2.552
25	121369	AA405657	Hs.128791 Human DNA sequence from clone 967N21 on chromosome 20p12.3-13. Contains	2.55
	114860	AA235112	Hs.106227 ESTs; Moderately similar to similar to murine RNA-binding protein [H.sapiens]	2.549
	121857	AA426017	Hs.62694 ESTs; Highly similar to DNA-REPAIR PROTEIN COMPLEMENTING	2.548
	110190	H20560	Hs.244624 ESTs	2.548
	132573	AA045333	Hs.51743 ESTs; Weakly similar to !! ALU SUBFAMILY SB2 WARNING ENTRY !! [H.sapiens]	2.542
	109706	F09729	Hs.12780 ESTs	2.537
30	135109	AA410391	Hs.94592 klotho	2.525
	132810	R37027	Hs.5737 KIAA0475 gene product	2.525
	124879	R73588	Hs.101533 ESTs	2.525
	103840	AA174190	Hs.50932 ESTs	2.525
35	119066	R22196	Hs.34492 ESTs	2.519
	114833	AA234362	Hs.87310 ESTs; Moderately similar to CGI-66 protein [H.sapiens]	2.507
	112998	T23555	Hs.103288 ESTs	2.5
	123312	AA496258	Hs.99601 ESTs	2.499
	121873	AA426270	Hs.145696 splicing factor (CC1.3)	2.491
	123321	AA496884	Hs.23972 ESTs	2.491
40	107760	AA018042	Hs.95078 EST	2.483
	102580	U60808	Hs.152981 CDP-diacylglycerol synthase (phosphatidate cytidyltransferase) 1	2.481
	103053	X56741	Hs.5947 mel transforming oncogene (derived from cell line NK14)- RAB8 homolog	2.475
	124756	R38100	Hs.106294 ESTs	2.475
45	112936	T15665	Hs.6185 ESTs; Weakly similar to BcDNA.GH12174 [D.melanogaster]	2.475
	125178	W58202	Hs.125731 ESTs	2.475
	112423	R62447	Hs.22123 ESTs	2.471
	123515	AA600323	Hs.112535 EST	2.462
	102842	U95020	Hs.21903 calcium channel; voltage-dependent; beta 4 subunit	2.457
50	102400	U42390	Hs.171957 triple functional domain (PTPRF interacting)	2.455
	113187	T56056	Hs.9992 ESTs	2.452
	131687	L11066	Hs.3069 heat shock 70kD protein 9B (mortalin-2)	2.448
	115314	AA280583	Hs.256501 ESTs	2.437
	128211	AI206427	Hs.166707 ESTs; Highly similar to Ran-binding protein 2 [H.sapiens]	2.43
55	134281	L11005	Hs.81047 aldehyde oxidase 1	2.425
	115985	AA447709	Hs.132094 ESTs; Moderately similar to putative transcription factor CA150 [H.sapiens]	2.425
	111348	N90041	Hs.9585 ESTs	2.418
	129430	AA258842	Hs.197877 Homo sapiens clone 23777 putative transmembrane GTPase mRNA; partial cds	2.418
	133863	C13990	Hs.76930 synuclein; alpha (non A4 component of amyloid precursor)	2.417
60	111164	N66857	Hs.14808 ESTs; Weakly similar to !! ALU CLASS C WARNING ENTRY !! [H.sapiens]	2.416
	132143	AA257056	Hs.7972 KIAA0871 protein	2.412
	130330	M55047	Hs.154679 synaptotagmin 1	2.408
	114219	Z39451	Hs.27389 ESTs	2.406
	117101	H94043	Hs.24341 DKFZP586I1419 protein	2.403
65	125433	AA034325	Hs.54320 ESTs	2.4
	111099	N62506	Hs.21958 ESTs	2.4
	120323	AA195405	Hs.110347 Homo sapiens mRNA for alpha integrin binding protein 80; partial	2.397
	118624	N69998	Hs.21801 ESTs	2.394
	123570	AA608955	Hs.109653 ESTs	2.389
	123562	AA608893	Hs.190065 ESTs	2.388

	131546	AA262821	Hs.28578	muscleblind (Drosophila)-like	2.385
	103143	X66141	Hs.75535	myosin; light polypeptide 2; regulatory; cardiac; slow	2.384
	123645	AA609310	Hs.188691	ESTs	2.383
5	130123	AA001835	Hs.150390	zinc finger protein 262	2.379
	131682	AA428368	Hs.30654	ESTs	2.378
	115909	AA436666	Hs.59761	ESTs	2.375
	125168	W45574	Hs.252497	ESTs	2.372
	123973	C14805	Hs.182151	ESTs	2.361
10	135197	U76456		Homo sapiens tissue inhibitor of metalloproteinase 4 mRNA, complete cds	2.357
	118689	N71545	Hs.184544	ESTs	2.357
	107734	AA016225	Hs.93386	ESTs	2.354
	124590	N69220	Hs.41381	ESTs; Weakly similar to ubiquitin hydrolyzing enzyme I [H.sapiens]	2.35
	111163	N66850	Hs.17606	ESTs	2.348
	112349	R58877	Hs.22665	ESTs; Moderately similar to dJ83L6.1 [H.sapiens]	2.345
15	129076	AA262179	Hs.169343	ESTs	2.345
	134238	R81509	Hs.184571	splicing factor; arginine/serine-rich 11	2.341
	116766	H13260	Hs.95097	ESTs	2.336
	106331	AA436853	Hs.34795	ESTs	2.333
	129003	AA443752	Hs.10784	ESTs	2.332
20	132368	AA599814	Hs.46637	ESTs; Weakly similar to cDNA EST yk289g5.5 comes from this gene [C.elegans]	2.332
	124697	R06273	Hs.186467	ESTs; Modly smlr to !! ALU SUBFAMILY J WARNING ENTRY !! [H.sapiens]	2.322
	120273	AA176688	Hs.221139	ESTs	2.313
	127110	AA304993	Hs.100861	ESTs; Weakly similar to p60 katanin [H.sapiens]	2.307
	105450	AA252621	Hs.93842	ESTs	2.301
25	119819	W74371	Hs.58383	ESTs	2.297
	102302	U33052	Hs.69171	protein kinase C-like 2	2.288
	130596	N74353	Hs.16475	ESTs	2.282
	114161	Z38904	Hs.22385	ESTs; Weakly similar to KIAA0970 protein [H.sapiens]	2.278
	130542	U64675		Human sperm membrane protein BS-63 mRNA, complete cds	2.277
30	104491	N71513	Hs.39328	ESTs	2.275
	116988	H82527		ys69e12.s1 Soares retina N2b4HR Homo sapiens cDNA clone	2.275
	126823	AA370120	Hs.7870	ESTs; Weakly similar to Ylr350wp [S.cerevisiae]	2.273
	108800	AA129731	Hs.90424	ESTs	2.273
35	101310	L41607	Hs.934	glucosaminyl (N-acetyl) transferase 2; l-branching enzyme	2.269
	126842	W19498	Hs.21085	ESTs	2.255
	127251	AA936428	Hs.128638	ESTs	2.251
	124647	N91947	Hs.125033	ESTs	2.249
	127112	AI143906	Hs.125103	ESTs	2.247
40	101973	S82597	Hs.80120	UDP-N-acetyl-alpha-D-galactosamine:polypeptide	2.246
	120999	AA398302	Hs.127437	ESTs	2.245
	130225	AA599583	Hs.15299	HMBA-inducible	2.243
	119980	W88678	Hs.249247	heterogeneous nuclear protein similar to rat helix destabilizing protein	2.243
	124222	H61053	Hs.222844	ESTs	2.24
	129199	H90914	Hs.128629	ESTs	2.236
45	106802	AA479101	Hs.16570	ESTs; Weakly similar to !! ALU SUBFAMILY SQ WARNING ENTRY !! [H.sapiens]	2.231
	126160	N90960	Hs.247277	ESTs; Weakly similar to transformation-related protein [H.sapiens]	2.229
	104627	AA001976	Hs.19603	ESTs	2.228
	106474	AA450212	Hs.42484	Homo sapiens mRNA; cDNA DKFZp564C053 (from clone DKFZp564C053)	2.226
	113096	T40927	Hs.8345	ESTs	2.225
50	135336	AA452822	Hs.99027	ESTs	2.225
	135344	R62976	Hs.168491	ESTs; Moderately similar to TRF1-interacting ankyrin-related	2.225
	126156	AA508354	Hs.118448	ESTs; Moderately similar to AKT3 protein kinase [H.sapiens]	2.222
	128885	AA397841	Hs.180141	cofilin 2 (muscle)	2.218
	107900	AA026385	Hs.176600	ESTs; Moderately similar to !! ALU SUBFAMILY SB2 WARNING	2.217
55	114481	AA033562	Hs.151572	ESTs	2.212
	109292	AA199828	Hs.188662	ESTs	2.212
	104257	AF006265	Hs.9222	estrogen receptor-binding fragment-associated gene 9	2.209
	132932	T15482	Hs.6093	ESTs	2.204
	127392	AA262728	Hs.14896	Homo sapiens clone 24590 mRNA sequence	2.204
60	104641	AA004652	Hs.18564	ESTs	2.2
	122529	AA449828	Hs.99229	ESTs	2.195
	124307	H93562	Hs.162395	proline synthetase co-transcribed (bacterial homolog)	2.193
	133601	S95936	Hs.75155	transferrin	2.193
	119904	W85709	Hs.128927	ESTs; Weakly similar to !! ALU SUBFAMILY SP WARNING ENTRY !! [H.sapiens]	2.192
65	100348	D64109	Hs.4994	transducer of ERBB2; 2 (TOB2)	2.185
	126871	AA351779	Hs.200334	ESTs	2.18
	127793	AI298835	Hs.30445	ESTs; Weakly similar to transcription regulator Staf-50 [H.sapiens]	2.178
	105149	AA169253	Hs.8958	ESTs	2.177
	121367	AA405648		zw39g8.s1 Soares_total_fetus_Nb2HF8_9w H sapiens cDNA clone IMAGE:772478	2.177

	111836	R36228	Hs.25119	ESTs	2.175
	133394	R16759	Hs.237225	ribosomal protein S5 pseudogene 1	2.175
	123207	AA489697	Hs.145053	ESTs	2.175
	129801	F11087	Hs.239666	ESTs	2.175
5	103393	X94612	Hs.41749	protein kinase; cGMP-dependent; type II	2.161
	132415	AA043223	Hs.4815	nudix (nucleoside diphosphate linked moiety X)-type motif 3	2.157
	106369	AA443828	Hs.25324	ESTs	2.157
	122963	AA478446	Hs.69559	KIAA1096 protein	2.156
	133473	M19309	Hs.73980	troponin T1; skeletal; slow	2.155
10	134257	C06270	Hs.8078	Homo sapiens mRNA; cDNA DKFZp586L081 (from clone DKFZp586L081)	2.155
	135156	AA056012	Hs.9552	binder of Arl Two	2.151
	104055	AA393755	Hs.117211	ESTs; Highly similar to CGI-62 protein [H.sapiens]	2.15
	102313	U33921	HSU33921	Clontech adult lung cDNA library (HL1158a) Homo sapiens cDNA	2.15
	109788	F10638	Hs.12432	Homo sapiens clone 24407 mRNA sequence	2.15
15	103507	Y10032	Hs.159640	serum/glucocorticoid regulated kinase	2.15
	116000	AA448710	Hs.41327	ESTs	2.15
	105858	AA399164	Hs.227676	ESTs; Moderately similar to !! ALU SUBFAMILY SQ	2.137
	103153	X66534	Hs.75295	guanylate cyclase 1; soluble; alpha 3	2.137
	126202	AA652238	Hs.199726	ESTs	2.135
20	115955	AA446121	Hs.44198	Homo sapiens BAC clone RG054D04 from 7q31	2.134
	104164	AA458770	Hs.27023	KIAA0917 protein	2.132
	108692	AA121270	Hs.82960	ESTs	2.128
	122878	AA465341	Hs.99640	ESTs	2.126
	134771	L13939	Hs.89576	adaptor-related protein complex 1; beta 1 subunit	2.125
25	104298	D31120	Hs.40368	adaptor-related protein complex 1; sigma 2 subunit	2.125
	104840	AA039595	Hs.42458	Homo sapiens mRNA; cDNA DKFZp586C1817 (from clone DKFZp586C1817)	2.125
	122180	AA435798	Hs.98835	ESTs; Moderately similar to putative ring zinc finger protein	2.125
	131012	H01992	Hs.202949	KIAA1102 protein	2.125
	134092	H17490	Hs.7905	ESTs; Highly similar to sorting nexin 9 [H.sapiens]	2.123
30	118617	N69666	Hs.183413	ESTs; Modtly smlr to !! ALU SUBFAMILY J WARNING ENTRY !! [H.sapiens]	2.123
	107155	AA621202	Hs.7946	DKFZP586D1519 protein	2.12
	130925	N71935	Hs.169378	multiple PDZ domain protein	2.12
	135167	U63717	Hs.95821	osteoclast stimulating factor 1	2.118
	105952	AA405263	Hs.181400	ESTs	2.109
35	110308	H38148	Hs.32775	ESTs	2.108
	116368	AA521186	Hs.94217	ESTs	2.107
	132939	U76189	Hs.61152	exostoses (multiple)-like 2	2.102
	117881	N50073	Hs.84926	ESTs; Highly similar to B-IND1 protein [M.musculus]	2.1
	112723	AA419622	Hs.104800	ESTs; Weakly similar to Mouse 19.5 mRNA; complete cds [M.musculus]	2.096
40	103500	Y09443	Hs.22580	alkylglycerone phosphate synthase	2.094
	121429	AA406293	Hs.193498	ESTs	2.093
	134632	AA398710	Hs.174139	chloride channel 3	2.091
	129785	F10980	Hs.184780	ESTs	2.09
	111065	N58193	Hs.18740	ESTs; Weakly similar to 1-evidence	2.089
45	114710	AA129931	Hs.79081	protein phosphatase 1; catalytic subunit; gamma isoform	2.083
	132711	N73702	Hs.238927	ESTs	2.083
	133377	R05490	Hs.7239	SEC24 (S. cerevisiae) related gene family; member B	2.079
	124773	R40923	Hs.106604	ESTs	2.078
	117759	N47587	Hs.97345	ESTs; Weakly similar to TROPOMODULIN [H.sapiens]	2.076
50	127386	AI457411	Hs.106728	ESTs	2.076
	101167	L15309	Hs.193677	zinc finger protein 141 (clone pHZ-44)	2.075
	109597	F02582	Hs.14474	ESTs	2.074
	124390	N29325	Hs.7535	ESTs; Highly similar to COBW-like placental protein [H.sapiens]	2.07
	116225	AA478609	Hs.47278	Human Chromosome 16 BAC clone CIT987SK-A-735G6	2.07
55	131243	R16667	Hs.24752	spectrin SH3 domain binding protein 1	2.069
	130557	T90830	Hs.15981	ESTs; Weakly similar to line-1 protein ORF2 [H.sapiens]	2.067
	134103	D14826	Hs.155924	cAMP responsive element modulator	2.064
	108833	AA131866	Hs.61661	ESTs; Weakly similar to DY3.6 [C.elegans]	2.063
	112286	R53765	Hs.158135	KIAA0981 protein	2.063
60	125624	AA165411	zq49a01.r1	Stratagene hNT neuron (#937233) Homo sapiens cDNA clone	2.061
	124612	N72200	Hs.13913	ESTs	2.058
	116335	AA495830	Hs.87013	ESTs	2.057
	112248	R51361	Hs.23423	ESTs	2.056
	115789	AA424754	Hs.43149	ESTs	2.056
65	107029	AA599219	Hs.187492	ESTs; Weakly similar to ALR [H.sapiens]	2.056
	110294	H30270	Hs.165062	ESTs	2.054
	120532	AA262354	Hs.186648	ESTs	2.054
	118180	N59249	Hs.48349	ESTs	2.052
	132018	AA293194	Hs.3737	ESTs	2.052

	132617	AA171913	Hs.5338	carbonic anhydrase XII	2.05
	131526	N36167	Hs.28274	ESTs	2.05
	113254	T64438	Hs.11449	DKFZP564O123 protein	2.05
5	122785	AA459978	Hs.99508	ESTs	2.05
	107203	D20426	Hs.5656	EST	2.05
	105713	AA291321	Hs.184319	ESTs; Moderately similar to KIAA1006 protein [H.sapiens]	2.046
	129385	D82675	Hs.110950	Homo sapiens clone 25007 mRNA sequence	2.042
	119116	R43845	Hs.64595	DKFZP566E2346 protein	2.04
10	116405	AA600253	Hs.55601	ESTs; Highly similar to host cell factor 2 [H.sapiens]	2.04
	125924	AA526849	Hs.82109	syndecan 1	2.039
	105599	AA279442	Hs.143460	protein kinase C; nu	2.037
	119741	W70205	Hs.43670	kinesin family member 3A	2.037
	101449	M21494	Hs.118843	creatine kinase; muscle	2.036
15	107109	AA609943	Hs.32793	ESTs	2.034
	117040	H89112		yw25e5.s1 Morton Fetal Cochlea Homo sapiens cDNA clone IMAGE:25328	2.034
	132906	AA142857	Hs.234896	ESTs; Highly similar to geminin [H.sapiens]	2.031
	105479	AA255546	Hs.23467	ESTs	2.027
	102031	U04898	Hs.2156	RAR-related orphan receptor A	2.027
20	119846	W80363	Hs.58446	ESTs	2.024
	124809	R46482	Hs.106875	ESTs	2.024
	130286	AA041548	Hs.154023	KIAA0573 protein	2.023
	124457	N50114	Hs.128704	ESTs	2.017
	125144	W37999	Hs.24336	ESTs	2.017
25	120581	AA281257	Hs.125868	ESTs	2.014
	104931	AA062731	Hs.108319	thyroid hormone receptor-associated protein; 150 kDa subunit	2.012
	120548	AA278846	Hs.187634	ESTs	2.011
	113933	W81362	Hs.30567	ESTs	2.011
	123072	AA485041	Hs.104308	ESTs	2.009
30	123648	AA609323	Hs.112689	ESTs	2.008
	116875	H67749	Hs.161022	EST	2.003
	103179	X69398	Hs.82685	CD47 antigen (Rh-related antigen; integrin-associated signal transducer)	1.995
	103478	Y07755	Hs.38991	S100 calcium-binding protein A2	1.995
	111007	N53378	Hs.22543	ESTs	1.995
35	120470	AA251797		zs11f3.s1 NCL_CGAP_GCB1 Homo sapiens cDNA clone	1.989
	112280	R53457	Hs.26040	ESTs; Weakly similar to fatty acid omega-hydroxylase [H.sapiens]	1.989
	114127	Z38652	Hs.106961	ESTs; Weakly similar to TYL [H.sapiens]	1.988
	129863	AA151005	Hs.129872	sperm surface protein	1.988
	106320	AA436608		ESTs	1.988
40	108933	AA147224	Hs.71814	ESTs	1.986
	105906	AA401633	Hs.22380	ESTs	1.982
	109029	AA157911	Hs.72200	ESTs	1.982
	118470	N66769	Hs.82781	ESTs	1.975
	115358	AA281886	Hs.88923	ESTs	1.975
45	115257	AA279060	Hs.193516	B-cell CLL/lymphoma 10	1.974
	126879	AA719776		zh38g04.s1 Soares_pineal_gland_N3HPG Homo sapiens cDNA clone IMAGE:414390	1.974
	109547	F01479	Hs.26966	ESTs	1.973
	127111	AA805726	Hs.220509	ESTs	1.969
	101266	L36645	Hs.73964	EphA4	1.966
50	129319	AA037467	Hs.30340	ESTs	1.965
	106211	AA428240	Hs.126083	ESTs	1.962
	112753	R93696	Hs.169882	ESTs	1.961
	120489	AA255538	Hs.190504	ESTs	1.959
	129699	AA458578	Hs.12017	KIAA0439 protein; homolog of yeast ubiquitin-protein ligase Rsp5	1.956
55	105425	AA251129	Hs.24416	ESTs	1.953
	134740	L37362	Hs.89455	opioid receptor; kappa 1	1.95
	109324	AA210700	Hs.86405	Homo sapiens mRNA; cDNA DKFZp564P056 (from clone DKFZp564P056)	1.95
	124303	H93043	Hs.107070	ESTs	1.95
	102337	U36922		Human fork head domain protein (FKHR) mRNA, 3' end	1.948
60	109441	AA228100	Hs.86998	nuclear factor of activated T-cells 5	1.946
	127364	AA179573	Hs.90061	progesterone binding protein	1.942
	105255	AA227498	Hs.3623	ESTs	1.942
	130672	L19783	Hs.177	phosphatidylinositol glycan; class H	1.942
	104301	D45332	Hs.6783	ESTs	1.94
65	132442	R62589	Hs.167419	ESTs	1.939
	105519	AA258063	Hs.23438	ESTs	1.937
	132902	AA490969	Hs.168147	ESTs	1.936
	118873	N89881	Hs.44577	ESTs	1.936
	114124	Z38595	Hs.125019	ESTs; Highly similar to KIAA0886 protein [H.sapiens]	1.934
	115075	AA255486	Hs.88045	ESTs	1.933

	110695	H93463	Hs.124777	ESTs	1.931
	105360	AA236209	Hs.187626	ESTs	1.931
	124998	T56013	Hs.77910	3-hydroxy-3-methylglutaryl-Coenzyme A synthase 1 (soluble)	1.929
	121816	AA424814	Hs.187509	ESTs	1.927
5	111717	R23241	Hs.110776	STAT induced STAT inhibitor-2	1.925
	128874	H06245	Hs.106801	ESTs	1.925
	109391	AA219699	Hs.184245	KIAA0929 protein Msx2 interacting nuclear target (MINT) homolog	1.913
	126129	H82165	Hs.40334	ESTs	1.911
	115553	AA369027	Hs.71414	ESTs	1.905
10	113811	W44928	Hs.4878	ESTs	1.905
	108345	AA070906		zm66d1.s1 Stratagene neuroepithelium (#937231) Homo sapiens cDNA clone	1.904
	120472	AA251875	Hs.104472	ESTs; Weakly similar to Gag-Pol polyprotein [M.musculus]	1.903
	116602	D80063	Hs.241673	EST	1.901
	121121	AA399371	Hs.189095	ESTs; Weakly similar to zinc finger protein SALL1 [H.sapiens]	1.9
15	125330	AA401804	Hs.114574	ESTs	1.896
	130095	F01831	Hs.14838	ESTs	1.894
	119782	W72982	Hs.58262	ESTs	1.894
	104115	AA428090	Hs.26102	ESTs	1.893
20	131313	C17938	Hs.22370	Homo sapiens mRNA; cDNA DKFZp564O0122 (from clone DKFZp564O0122)	1.891
	105583	AA278907	Hs.24549	ESTs	1.891
	122825	AA461195	Hs.99580	ESTs	1.887
	119495	W35390	Hs.55533	ESTs	1.886
	130309	AA134289	Hs.15423	Homo sapiens BAC clone RG114B19 from 7q31.1	1.886
	125628	AA418069	Hs.241493	natural killer-tumor recognition sequence	1.886
25	110611	H66947	Hs.14671	ESTs; Highly similar to gene ERCC5 protein [H.sapiens]	1.885
	117301	N22569	Hs.43215	ESTs	1.884
	131406	N92239	Hs.26471	Wnt inhibitory factor-1	1.881
	126428	AA013312	Hs.64988	ESTs	1.881
	120285	AA182882	Hs.111110	titin-cap (telethonin)	1.878
30	112724	R91753	Hs.17757	ESTs	1.878
	103121	X63679	Hs.4147	translocating chain-associating membrane protein	1.875
	124381	N26765	Hs.109008	ESTs	1.875
	117226	N20468	Hs.177322	ESTs; Weakly similar to putative p150 [H.sapiens]	1.875
	105610	AA279991	Hs.124691	ESTs; Weakly similar to trithorax homologue 2 [H.sapiens]	1.875
35	111229	N69113	Hs.110855	ESTs	1.875
	120627	AA285079	Hs.190474	ESTs	1.873
	107048	AA600012	Hs.10669	ESTs; Moderately similar to KIAA0400 [H.sapiens]	1.872
	104041	AA381902	Hs.197114	RNA binding protein	1.872
	115162	AA258366	Hs.227806	ras GTPase activating protein-like	1.872
40	102239	U26726	Hs.1376	hydroxysteroid (11-beta) dehydrogenase 2	1.87
	100043	M10098		AFFX control: 18S ribosomal RNA	1.868
	120296	AA191353	Hs.22385	ESTs; Weakly similar to KIAA0970 protein [H.sapiens]	1.867
	129011	S72869	Hs.107932	DNA segment; single copy; probe pH4 (transforming sequence; thyroid-1;	1.867
	134851	R44479	Hs.90232	KIAA0552 gene product	1.866
45	117392	N26175	Hs.93405	ESTs	1.864
	114530	AA053027	Hs.191797	ESTs	1.863
	123541	AA608794	Hs.112592	ESTs	1.863
	124890	R78618	Hs.34145	ESTs; Weakly similar to RAS-RELATED PROTEIN RAB-8 [H.sapiens]	1.862
	105299	AA233511	Hs.194720	ATP-binding cassette; sub-family G (WHITE); member 2	1.861
50	103560	Z20656	Hs.182787	myosin; heavy polypept 6; cardiac muscle; alpha (cardiomyopathy; hypertrophic 1)	1.861
	113073	T33637	Hs.6841	ESTs	1.86
	120407	AA235040	Hs.107283	ESTs	1.859
	103892	AA243523	Hs.17155	ESTs	1.858
	123795	AA620381	Hs.70488	ESTs	1.857
55	108524	AA084323	Hs.68138	ESTs	1.857
	113953	W85812	Hs.187554	ESTs	1.856
	110721	H97678	Hs.31319	ESTs	1.856
	129426	AA412087	Hs.168272	EST; Highly smlr to prot inhibitor of activated STAT prot PIASx-alpha [H.sapiens]	1.853
	112102	R44840	Hs.21303	ESTs	1.852
60	118502	N67317	Hs.50150	ESTs	1.852
	107619	AA004955	Hs.60015	ESTs	1.851
	100436	D87446	Hs.75912	KIAA0257 protein	1.85
	120652	AA287312	Hs.191648	ESTs	1.85
	121643	AA417078	Hs.193767	ESTs	1.843
65	117387	N26011	Hs.53810	ESTs	1.843
	132084	Y12394	Hs.3886	karyopherin alpha 3 (importin alpha 4)	1.843
	124449	N48593	Hs.121820	ESTs	1.841
	120263	AA173440	Hs.193919	ESTs	1.838
	127226	AA731036	Hs.3463	ribosomal protein S23	1.838

	111837	R36447	Hs.24453	ESTs	1.835
	128727	M64174	Hs.50651	Janus kinase 1 (a protein tyrosine kinase)	1.834
	114439	AA018937	Hs.128629	ESTs	1.833
5	102332	U35637		Human nebulin mRNA, partial cds	1.83
	126579	W72979	Hs.146082	ESTs	1.83
	102341	U37122	Hs.8110	adducin 3 (gamma)	1.83
	114246	Z39848	Hs.12079	ESTs	1.828
	131757	D17532	Hs.316	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 6 (RNA helicase; 54kD)	1.823
10	108904	AA136521	Hs.71148	ESTs; Weakly similar to putative p150 [H.sapiens]	1.823
	115084	AA255566	Hs.42484	Homo sapiens mRNA; cDNA DKFZp564C053 (from clone DKFZp564C053)	1.823
	131957	AA609008	Hs.183232	ESTs	1.822
	100131	D12485	Hs.11951	phosphodiesterase I/nucleotide pyrophosphatase 1 (homologous to mouse Ly-41 antigen)	1.822
15	124163	H30539	Hs.189838	ESTs	1.821
	118204	N59859	Hs.48443	ESTs	1.821
	107727	AA016021	Hs.173091	DKFZP434K151 protein	1.82
	100357	D78156	Hs.241548	RAS p21 protein activator 2	1.82
	116295	AA489016	Hs.91216	ESTs; Highly similar to partial CDS; human putative tumor suppressor [H.sapiens]	1.82
20	124833	R54112	Hs.128697	ESTs	1.817
	122587	AA453255	Hs.6968	ESTs	1.817
	114359	Z41589	Hs.153483	ESTs; Moderately similar to H1 chloride channel [H.sapiens]	1.815
	111289	N72253	Hs.238246	ESTs	1.813
	110826	N30068	Hs.15347	ESTs	1.812
25	104106	AA422123	Hs.42457	ESTs	1.811
	130043	AA055404	Hs.193953	ESTs; Weakly similar to !! ALU SUBFAMILY J WARNING ENTRY !! [H.sapiens]	1.253
	115864	AA432080	Hs.81200	ESTs	1.81
	129737	AA056140	Hs.122684	ESTs	1.81
	124477	N53158	Hs.102682	ESTs	1.809
30	100782	HG3740-HT4010		Basic Transcription Factor 2, 34 Kda Subunit	1.806
	106101	AA421053	Hs.34395	ESTs	1.806
	115479	AA287596		zs52h09.s1 NCI_CGAP_GCB1 H sapiens cDNA clone IMAGE:701153	1.804
	116104	AA456635	Hs.78524	ESTs	1.804
	114173	Z39050	Hs.21963	ESTs	1.804
35	132632	N59764	Hs.5398	guanine-monophosphate synthetase	1.803
	119135	R49548	Hs.169681	death effector domain-containing	1.802
	131559	N91087	Hs.28728	ESTs; Weakly similar to F55A12.9 [C.elegans]	1.801
	126922	AA177138	Hs.161671	ESTs	1.8
	117375	N25427	Hs.108812	ESTs	1.8
40	103571	Z25535	Hs.211608	nucleoporin 153kD	1.8
	105978	AA406367	Hs.15973	ESTs	1.8
	125904	H22372	Hs.163586	ESTs	1.799
	133883	AA397915	Hs.77221	choline kinase	1.798
	105777	AA348412	Hs.23096	ESTs	1.797
45	110166	H19480	Hs.174309	ESTs	1.796
	105038	AA130273	Hs.7584	ESTs; Weakly similar to hypothetical protein; similar to [H.sapiens]	1.796
	105427	AA251330	Hs.28248	ESTs	1.795
	115278	AA279757	Hs.67466	ESTs; Weakly similar to BACN32G11.d [D.melanogaster]	1.794
	133104	L13698	Hs.65029	growth arrest-specific 1	1.794
	131170	N48674	Hs.23796	Human DNA sequence from clone 1052M9 on chromosome Xq25. Contains the	1.792
50	100136	D13540	Hs.22868	protein tyrosine phosphatase; non-receptor type 11	1.791
	127263	AA331157		EST35035 Embryo, 6 week, subtracted (total cDNA) I Homo sapiens cDNA	1.79
	114157	Z38878	Hs.24979	ESTs	1.79
	125601	AI096717	Hs.247043	KIAA0525 protein	1.788
55	118472	N66818	Hs.42179	ESTs	1.787
	112456	R63925	Hs.28464	ESTs	1.787
	130236	N69682	Hs.51957	SC35-interacting protein 1	1.786
	133297	AA600057	Hs.70266	KIAA0905 protein	1.784
	125650	R40096	Hs.176578	ESTs	1.784
	132056	T89386	Hs.38176	KIAA0606 protein; SCN Circadian Oscillatory Protein (SCOP)	1.783
60	129093	AA262710	Hs.108614	KIAA0627 protein	1.783
	123176	AA489020	Hs.193424	ESTs	1.782
	106340	AA441792	Hs.22857	chord domain-containing protein 1	1.781
	100598	HG2463-HT2559		Guanine Nucleotide-Binding Protein G25k	1.779
65	104038	AA374532		EST86676 HSC172 cells I Homo sapiens cDNA 5' end, mRNA sequence	1.778
	122235	AA436475	Hs.190104	ESTs	1.777
	105104	AA151771	Hs.76941	ATPase; Na+/K+ transporting; beta 3 polypeptide	1.776
	107601	AA004636	Hs.50223	ESTs	1.776
	131467	W68255	Hs.27194	DKFZP434K171 protein	1.776
	118449	N66413	Hs.172466	ESTs; Weakly similar to KIAA0775 protein [H.sapiens]	1.776

	107969	AA034030	Hs.155212	methylmalonyl Coenzyme A mutase	1.775
	115527	AA342079	Hs.252055	ESTs	1.775
	132471	T16305	Hs.49349	beta-site APP-cleaving enzyme	1.775
5	105966	AA406105	Hs.5344	adaptor-related protein complex 1; gamma 1 subunit	1.774
	127548	AA373091	Hs.93832	Homo sapiens clone 24483 unknown mRNA; parital cds	1.774
	106217	AA428379	Hs.24870	ESTs	1.773
	131214	N26777	Hs.172635	ESTs	1.773
	106295	AA435664	Hs.8583	similar to APOBEC1	1.773
10	106328	AA436705	Hs.28020	KIAA0766 gene product	1.772
	124661	N93797	Hs.3090	EphB1	1.772
	122988	AA479166	Hs.105633	ESTs	1.772
	115504	AA291946	Hs.42736	ESTs	1.771
	105168	AA180208	Hs.16606	ESTs; Highly similar to CGI-32 protein [H.sapiens]	1.767
	129153	AA188618	Hs.181461	ariadne; Drosophila; homolog of	1.766
15	105829	AA398290	Hs.21965	ESTs	1.764
	101811	M86917	Hs.24734	oxysterol binding protein	1.764
	100138	D13628	Hs.2463	angiopoietin 1	1.764
	124704	R07335		ye96c1.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone	1.763
20	122314	AA442257	Hs.192076	ESTs	1.762
	109865	H02566	Hs.191268	Homo sapiens mRNA; cDNA DKFZp434N174 (from clone DKFZp434N174)	1.761
	106206	AA428069	Hs.89519	KIAA1046 protein	1.758
	107135	AA620782	Hs.23247	ESTs	1.757
	105760	AA338960	Hs.28170	ESTs	1.756
25	106288	AA435536	Hs.24336	ESTs	1.756
	103968	AA304566	Hs.3542	ESTs	1.756
	129559	AA234945	Hs.11360	ESTs	1.756
	117885	N50112	Hs.47023	ESTs	1.754
	107032	AA599472	Hs.247309	succinate-CoA ligase; GDP-forming; beta subunit	1.754
30	124807	R45963	Hs.233811	ESTs; Weakly similar to ORF2 [M.musculus]	1.753
	100276	D42047	Hs.82432	KIAA0089 protein	1.753
	110924	N47938		yy84a09.s1 Soares_multiple_sclerosis_2NbHMSP Homo sapiens cDNA clone	1.751
	133002	AF006082	Hs.62461	ARP2 (actin-related protein 2; yeast) homolog	1.751
	132530	AA455917	Hs.50785	SEC22; vesicle trafficking protein (S. cerevisiae)-like 1	1.75
35	110759	N21671	Hs.19025	ESTs	1.75
	106138	AA424515	Hs.33264	ESTs	1.75
	107348	U43701	Hs.184776	ribosomal protein L23a	1.75
	115867	AA432162	Hs.165986	DKFZP586B2022 protein	1.749
	135398	AA194075	Hs.99908	nuclear receptor coactivator 4	1.747
40	113783	W19222	Hs.7041	ESTs; Weakly similar to !! ALU SUBFAMILY SQ WARNING ENTRY !! [H.sapiens]	1.747
	134898	X98330	Hs.90821	ryanodine receptor 2 (cardiac)	1.745
	132215	T10132	Hs.4236	KIAA0478 gene product	1.744
	104229	AB002346	Hs.61289	synaptojanin 2	1.743
	116166	AA461556	Hs.202949	KIAA1102 protein	1.743
45	115433	AA284252	Hs.58372	ESTs	1.743
	114908	AA236545	Hs.54973	ESTs	1.742
	127425	AA470941	Hs.143162	ESTs	1.741
	131089	Z38807	Hs.22870	ESTs	1.739
	113498	T88908	Hs.189746	ESTs	1.738
50	116710	F10577	Hs.70312	ESTs	1.735
	127210	R51476		yg76f04.r1 Soares infant brain 1NIB Homo sapiens cDNA clone	1.733
	120554	AA279654	Hs.194524	ESTs	1.733
	129940	U18242	Hs.13572	calcium modulating ligand	1.732
	117023	H88157	Hs.41105	ESTs	1.731
55	111700	R22212	Hs.23361	ESTs	1.731
	116911	H72240	Hs.39292	ESTs; Moderately similar to KIAA0745 protein [H.sapiens]	1.731
	106025	AA412063	Hs.6065	ESTs	1.728
	108626	AA101984	Hs.61697	G-protein coupled receptor	1.726
	111614	R12581	Hs.191146	ESTs	1.726
60	134134	L76703	Hs.173328	protein phosphatase 2; regulatory subunit B (B56); epsilon isoform	1.725
	106886	AA489086	Hs.36545	ESTs	1.725
	117998	N52136	Hs.93828	ESTs	1.725
	121204	AA400422	Hs.55896	ESTs	1.725
	121342	AA404995	Hs.192480	ESTs	1.725
65	131129	R27296	Hs.23240	ESTs	1.725
	116235	AA479181	Hs.186726	ESTs	1.725
	102423	U44754	Hs.179312	small nuclear RNA activating complex; polypeptide 1; 43kD	1.724
	110273	H29050	Hs.24096	ESTs	1.722
	108758	AA127395	Hs.222414	ESTs	1.722
	110672	H88477	Hs.191178	ESTs	1.721

	120271	AA176404	Hs.111092	ESTs; Weakly similar to ZINC FINGER PROTEIN 136 [H.sapiens]	1.72
	100227	D28915	Hs.82316	interferon-induced; hepatitis C-associated microtubular aggregate prot (44kD)	1.719
	129232	W69459	Hs.109655	sex comb on midleg (Drosophila)-like 1	1.719
5	134663	W73367	Hs.8750	ESTs	1.717
	104902	AA055475	Hs.104143	clathrin; light polypeptide (Lca)	1.717
	120582	AA281290	Hs.125287	ESTs; Weakly similar to BC331191_1 [H.sapiens]	1.717
	134891	F03517	Hs.90787	ESTs	1.716
	106219	AA428567	Hs.26613	Homo sapiens mRNA; cDNA DKFZp586F1323 (from clone DKFZp586F1323)	1.715
10	116372	AA521311	Hs.13854	ESTs	1.713
	107570	AA001870	Hs.237323	N-acetylglucosamine-phosphate mutase; DKFZP434B187 protein	1.713
	106198	AA427816	Hs.11803	ESTs	1.712
	125136	W31479	Hs.129051	ESTs	1.712
	104973	AA085676	Hs.6763	KIAA0942 protein	1.712
	128710	J04813	Hs.104117	cytochrome P450; subfamily IIIA (niphedipine oxidase); polypeptide 5	1.711
15	123994	D20899	Hs.107127	Homo sapiens mRNA; cDNA DKFZp564G022 (from clone DKFZp564G022)	1.711
	127871	AA766511	Hs.128848	ESTs	1.71
	116089	AA455933	Hs.41324	ESTs	1.709
	123337	AA504153	Hs.132797	ESTs; Weakly similar to ORF YGL050w [S.cerevisiae]	1.708
	123619	AA609200	Hs.162686	ESTs	1.708
20	104781	AA026617	Hs.21610	ESTs; Highly similar to BA11-associated protein 1 [H.sapiens]	1.707
	115114	AA256468	Hs.88148	ESTs	1.705
	117852	N49408	Hs.136102	KIAA0853 protein	1.705
	127644	T57570	Hs.77039	ribosomal protein S3A	1.704
	111359	N91273	Hs.27179	ESTs	1.702
25	131721	L36644	Hs.31092	EphA5	1.7
	132438	F08925	Hs.48610	ESTs	1.7
	132476	N67192	Hs.49476	Homo sapiens clone TUA8 Cri-du-chat region mRNA	1.7
	130990	F02488	Hs.21917	KIAA0768 protein	1.7
	128499	AA487503	Hs.100636	ESTs	1.698
30	120780	AA342337	Hs.241569	ESTs; Modtly smlr to !! ALU SUBFAMILY SQ WARNING ENTRY !! [H.sapiens]	1.697
	132920	L06133	Hs.606	ATPase; Cu++ transporting; alpha polypeptide (Menkes syndrome)	1.696
	135037	U77948	Hs.184122	general transcription factor II; i	1.696
	110024	H11297	Hs.31050	ESTs	1.695
	134415	AA329274	Hs.82911	protein tyrosine phosphatase type IVA; member 2	1.694
35	102223	U24685	Hs.148226	Human anti-B cell autoantibody IgM heavy chain variable V-D-J region (VH4) gene; clone E11; VH4-63 non-productive rearrangement	1.694
	126712	AA205862	Hs.7942	ESTs	1.694
	101507	M27492	Hs.82112	interleukin 1 receptor; type I	1.692
	106291	AA435551	Hs.30824	ESTs	1.691
40	116826	H58691	Hs.8215	ESTs; Weakly similar to double-stranded RNA-binding nuclear protein DRBP76 [H.sapiens]	1.69
	135339	D59269	Hs.127842	Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 783648	1.69
	118250	N62602		yz75b6.s1 Soares_multiple_sclerosis_2NbHMSP Homo sapiens cDNA clone IMAGE:288851 3' similar to contains Alu repetitive element; mRNA sequence	1.689
45	106470	AA450116	Hs.186180	ESTs	1.688
	108203	AA057678	Hs.63408	ESTs	1.687
	119748	W70313	Hs.126906	ESTs	1.686
	116576	D51228	Hs.79404	neuron-specific protein	1.683
	123035	AA481392	Hs.105166	ESTs	1.683
50	126668	AA011616	Hs.184086	ESTs	1.681
	101512	M28209	Hs.250716	RAB1; member RAS oncogene family	1.678
	102704	U76638	Hs.54089	BRCA1 associated RING domain 1	1.677
	126218	AA256386	Hs.13649	Novel human gene mapping to chromosome 13; similar-to rat RhoGAP	1.676
	111180	N67277	Hs.9403	ESTs	1.676
55	105937	AA404342	Hs.173531	ESTs	1.675
	114118	Z38520	Hs.175930	ESTs	1.675
	109203	AA190634	Hs.108787	endoplasmic reticulum membrane protein	1.675
	125245	W86608	Hs.7243	ubiquitin specific protease 24	1.675
	102906	X06956	Hs.75318	tubulin; alpha 1 (testis specific)	1.675
60	125914	AA262925	Hs.180034	cleavage stimulation factor; 3' pre-RNA; subunit 3; 77kD	1.674
	134294	U63289	Hs.81248	CUG triplet repeat; RNA-binding protein 1	1.674
	109742	F10108	Hs.183333	ESTs	1.673
	134674	D63876	Hs.87726	KIAA0154 protein	1.673
	104079	AA402937	Hs.103238	ESTs	1.671
65	107554	AA001386	Hs.59844	ESTs	1.671
	132439	AA243139	Hs.4863	Homo sapiens clone 25088 mRNA sequence	1.669
	124515	N58172	Hs.109370	ESTs	1.668
	124300	H92575	Hs.105959	ESTs; Weakly similar to !! ALU SUBFAMILY SQ WARNING ENTRY !! [H.sapiens]	1.668
	126809	AA743475	Hs.171693	ESTs	1.667

	106095	AA419547	Hs.11713	ESTs	1.664
	101754	M77142	Hs.239489	TIA1 cytotoxic granule-associated RNA-binding protein	1.663
	105188	AA192306	Hs.23926	ESTs	1.663
	113582	T91371	Hs.16824	EST	1.661
5	119559	W38197		Accession not listed in Genbank	1.661
	119961	W87535	Hs.59015	ring finger protein 9	1.657
	123255	AA490890	Hs.105273	ESTs	1.657
	111078	N59230	Hs.186574	ESTs	1.655
	113082	T40528	Hs.8246	ESTs	1.654
10	119589	W44692	Hs.124177	ESTs	1.652
	104308	D53639	Hs.77904	ribosomal protein S26	1.65
	103073	X59417	Hs.74077	proteasome (prosome; macropain) subunit; alpha type; 6	1.65
	124424	N35314	Hs.107265	ESTs	1.65
15	128890	AA096157	Hs.182364	ESTs; Weakly similar to 25 kDa trypsin inhibitor [H.sapiens]	1.65
	119400	T92767		ye27d06.s1 Stratagene lung (#937210) Homo sapiens cDNA clone IMAGE:118955 3', mRNA sequence.	1.65
	131631	AA486868	Hs.29802	slit (Drosophila) homolog 2	1.65
	118229	N62339	Hs.180532	heat shock 90kD protein 1; alpha	1.649
	118533	N67954	Hs.49413	ESTs	1.648
20	130666	AA476307	Hs.194035	KIAA0737 gene product	1.647
	103093	X60708	Hs.44926	dipeptidylpeptidase IV (CD26; adenosine deaminase complexing protein 2)	1.647
	128667	U69140	Hs.103419	fasciculation and elongation protein zeta 2 (zyglin II)	1.646
	112933	T15530	Hs.221439	ESTs	1.646
	114546	AA056263	Hs.132747	ESTs	1.645
25	126705	AA579377	Hs.180532	heat shock 90kD protein 1; alpha	1.644
	114399	AA007595	Hs.220937	ESTs	1.642
	118836	N79820	Hs.50854	ESTs	1.64
	100401	D85423		Homo sapiens mRNA for Cdc5, partial cds	1.64
	105681	AA284865	Hs.171228	KIAA1040 protein	1.639
30	132526	AA460128	Hs.5074	similar to S. pombe dim1+	1.639
	133809	AA034002	Hs.76359	catalase	1.639
	115968	AA447083	Hs.134522	ESTs	1.637
	116370	AA521256	Hs.236204	ESTs; Moderately similar to NUCLEAR PORE COMPLEX PROTEIN NUP107 [R.norvegicus]	1.631
35	109644	F04477	Hs.204802	ESTs; Moderately similar to GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE; LIVER [H.sapiens]	1.627
	103427	X97303		H.sapiens mRNA for Ptg-12 protein	1.627
	132186	T33888	Hs.221040	KIAA1038 protein	1.626
	131428	U17838	Hs.26719	PR domain containing 2; with ZNF domain	1.626
40	126638	AA649257	Hs.188602	ESTs	1.625
	114503	AA039568	Hs.188083	ESTs	1.625
	121242	AA400857	Hs.97509	EST	1.625
	122414	AA446885	Hs.99087	ESTs; Moderately similar to ZINC FINGER PROTEIN 141 [H.sapiens]	1.625
	110632	H72344	Hs.171635	ESTs	1.624
45	111389	N95837	Hs.169111	ESTs; Weakly similar to L82A [D.melanogaster]	1.624
	112449	R63802	Hs.124186	ring finger protein 2	1.623
	113070	T33464	Hs.6298	ESTs	1.622
	107229	D59284	Hs.34644	ESTs	1.618
	132710	W93726	Hs.55279	protease inhibitor 5 (maspin)	1.617
50	124664	N94814	Hs.33540	ESTs; Weakly similar to KIAA0765 protein [H.sapiens]	1.617
	130166	AA350690	Hs.151411	KIAA0916 protein	1.616
	125040	T78451	Hs.199961	ESTs	1.615
	132972	H39627	Hs.164967	ESTs; Weakly similar to !! ALU SUBFAMILY SB WARNING ENTRY !! [H.sapiens]	1.615
	115873	AA433916	Hs.90093	heat shock 70kD protein 4	1.611
55	120408	AA235045	Hs.190151	ESTs	1.61
	120934	AA383773	Hs.191500	ESTs	1.61
	115259	AA279071	Hs.13453	splicing factor 3b; subunit 1; 155kD	1.609
	134330	D20113	Hs.8185	ESTs; Highly similar to CGI-44 protein [H.sapiens]	1.607
	115117	AA256492	Hs.49007	poly(A) polymerase	1.606
60	125162	W44682	Hs.109896	ESTs	1.605
	103946	AA285246	Hs.111650	ESTs; Weakly similar to Prt1 homolog [H.sapiens]	1.604
	133389	AA166917	Hs.72639	ESTs	1.603
	115528	AA342301	Hs.53929	ESTs; Weakly similar to !! ALU CLASS B WARNING ENTRY !! [H.sapiens]	1.602
	129704	W81301	Hs.12064	ubiquitin specific protease 22	1.602
65	109313	AA206800	Hs.86276	ESTs; Moderately similar to zinc finger protein dp [H.sapiens]	1.601
	130457	U58091	Hs.155976	cullin 4B	1.6
	123076	AA485211	Hs.190046	ESTs	1.6
	115113	AA256460	Hs.44610	ESTs	1.6
	117731	N46433	Hs.46609	ESTs	1.6

5	123344	AA504338	Hs.171857	ESTs	1.599
	131798	X86098	Hs.3238	adenovirus 5 E1A binding protein	1.597
	125370	AA256743	Hs.151791	KIAA0092 gene product	1.596
	114918	AA236813	Hs.72324	ESTs; Highly similar to unknown [H.sapiens]	1.596
	114807	AA160805	Hs.199832	ESTs	1.596
10	105103	AA151593	Hs.10130	ESTs	1.594
	125004	T60120		yb68f02.s1 Stratagene ovary (#937217) Homo sapiens cDNA clone IMAGE:76347 3', mRNA sequence.	1.592
	105658	AA282914	Hs.10176	ESTs	1.589
	110455	H52172		yt85e8.s1 Soares_pineaL_gland_N3HPG Homo sapiens cDNA clone IMAGE:23111 3' similar to contains Alu repetitive element, mRNA sequence	1.589
	119780	W72967	Hs.191381	ESTs; Weakly similar to hypothetical protein [H.sapiens]	1.587
15	126983	AA211537		zn55d01.r1 Stratagene muscle 937209 Homo sapiens cDNA clone IMAGE:562081 5', mRNA sequence.	1.586
	134675	AA250745	Hs.87773	protein kinase; cAMP-dependent; catalytic; beta	1.584
	105431	AA252033	Hs.15036	ESTs; Weakly similar to !! ALU SUBFAMILY J WARNING ENTRY !! [H.sapiens]	1.584
	120187	Z40251	Hs.56974	ESTs	1.584
	115830	AA428137	Hs.86434	ESTs	1.581
20	135069	AA456311	Hs.93961	ESTs; Weakly similar to !! ALU CLASS A WARNING ENTRY !! [H.sapiens]	1.581
	122997	AA479295	Hs.106290	Kelch motif containing protein	1.581
	119707	W67569	Hs.44143	ESTs; Weakly similar to SNF2alpha protein [H.sapiens]	1.58
	131934	D80948	Hs.34922	ESTs	1.58
	106141	AA424558	Hs.9302	phosducin-like	1.58
25	115271	AA279422	Hs.5724	ESTs	1.579
	131468	R27598	Hs.27197	KIAA0797 protein	1.577
	131165	R98173	Hs.23763	Max-interacting protein	1.575
	117273	N21680	Hs.43047	ESTs	1.575
	101569	M33772	Hs.182421	troponin C2; fast	1.575
30	116127	AA459703	Hs.79070	v-myc avian myelocytomatosis viral oncogene homolog	1.575
	120022	W90625	Hs.58432	ESTs	1.575
	117512	N32157	Hs.82207	ESTs	1.574
	106511	AA452865	Hs.206713	UDP-Gal:betaGlcNAc beta 1;4- galactosyltransferase; polypeptide 2	1.573
	116415	AA609204	Hs.27973	KIAA0874 protein	1.573
35	127879	AA810215	Hs.189079	ESTs	1.571
	125211	W72798	Hs.103177	ESTs; Wkly smlr to cDNA EST EMBL:D32579 comes from this gene [C.elegans]	1.571
	114746	AA135638	Hs.223756	ESTs	1.571
	122698	AA456112	Hs.99410	ESTs	1.57
	116765	H12636	Hs.121585	ESTs; Weakly similar to reverse transcriptase [H.sapiens]	1.568
40	130895	AA609828	Hs.21015	ESTs; Highly similar to tetracycline transporter-like protein [M.musculus]	1.568
	114338	Z41366	Hs.40109	KIAA0872 protein	1.567
	111005	N53076	Hs.5996	ESTs	1.567
	128135	AA913491	Hs.189143	ESTs; Modrtly smlr to !! ALU SUBFAMILY J WARNING ENTRY !! [H.sapiens]	1.567
	112046	R43365	Hs.22273	ESTs	1.566
45	132160	AA281770	Hs.184081	seven in absentia (Drosophila) homolog 1	1.566
	111568	R10153	Hs.20561	ESTs	1.566
	127775	H04106	Hs.179902	ESTs; Weakly similar to NG22 [H.sapiens]	1.566
	115359	AA281936	Hs.88914	ESTs	1.566
	121845	AA425734	Hs.165066	ESTs; Weakly similar to hypothetical protein [H.sapiens]	1.565
50	127854	AA769520		ESTs; Weakly similar to REGULATOR OF MITOTIC SPINDLE ASSEMBLY 1 [H.sapiens]	1.564
	120287	AA187679	Hs.111114	ESTs	1.563
	114940	AA243012	Hs.75928	ESTs	1.562
	126716	AA031700	Hs.251962	ESTs	1.562
	134161	U97188	Hs.79440	IGF-II mRNA-binding protein 3	1.561
55	125390	H95094	Hs.75187	translocase of outer mitochondrial membrane 20 (yeast) homolog	1.561
	115334	AA281244	Hs.65300	ESTs	1.559
	113721	T97931	Hs.18190	EST	1.558
	114895	AA236177	Hs.76591	KIAA0887 protein	1.558
	119341	T62571	Hs.146388	microtubule-associated protein 7	1.558
60	108012	AA039616	Hs.61933	ESTs	1.558
	130335	AA156499	Hs.8454	protein kinase; cAMP-dependent; regulatory; type II; alpha	1.557
	134351	R82074	Hs.82109	syndecan 1	1.557
	133300	D51401	Hs.70333	ESTs	1.553
	106920	AA490899	Hs.24462	ESTs	1.553
65	118744	N74075	Hs.94293	EST	1.552
	126489	W20016	Hs.144228	ESTs; Weakly similar to ZINC FINGER PROTEIN 83 [H.sapiens]	1.55
	115913	AA436720	Hs.65487	ESTs	1.55
	107868	AA025234	Hs.61260	ESTs	1.55
	134520	N21407	Hs.257325	ESTs	1.55

	109703	F09684	Hs.24792	ESTs; Weakly similar to ORF YOR283w [S.cerevisiae]	1.55
	120288	AA187938	Hs.55189	ESTs; Weakly similar to F25B5.3 [C.elegans]	1.548
	106356	AA443277	Hs.31034	peroxisomal biogenesis factor 11A	1.548
5	129460	AA235627	Hs.11171	APG5 (autophagy 5; S. cerevisiae)-like	1.547
	133950	D11961	Hs.77823	ESTs	1.546
	128172	AI400862	Hs.142607	ESTs	1.546
	114162	Z38909	Hs.22265	ESTs	1.545
	101803	M86546	Hs.155691	pre-B-cell leukemia transcription factor 1	1.544
	113617	T93630	Hs.17207	ESTs	1.542
10	104896	AA054228	Hs.23165	ESTs	1.541
	114477	AA032013	Hs.144260	EST	1.54
	110731	H98653	Hs.188006	KIAA0878 protein	1.54
	130367	Z38501	Hs.8768	ESTs; Wkly smlr to !! ALU SUBFAMILY SQ WARNING ENTRY !! [H.sapiens]	1.538
	130539	L07044	Hs.250857	Homo sapiens calcium/calmodulin-dependent protein kinase II mRNA; partial cds	1.538
15	134921	W60186	Hs.169487	Kreisler (mouse) maf-related leucine zipper homolog	1.537
	130583	W24957	Hs.16281	ESTs; Moderately similar to similar to C.elegans protein encoded in cosmid T20D3 [H.sapiens]	1.537
	133723	AA088851	Hs.75744	S-adenosylmethionine decarboxylase 1	1.537
	106450	AA449469	Hs.11859	ESTs	1.536
20	104120	AA429838	Hs.89519	KIAA1046 protein	1.536
	100533	HG1879-HT1919		Ras-Like Protein Tc10	1.535
	130664	R09049	Hs.17625	ESTs	1.535
	127122	AA279153	Hs.190049	ESTs	1.535
	134264	T03391	Hs.8087	ESTs	1.535
25	132319	AA418662	Hs.44625	ESTs	1.535
	115465	AA286941	Hs.43691	ESTs	1.533
	125003	T59442	Hs.100445	ESTs	1.532
	102273	U30888	Hs.75981	ubiquitin specific protease 14 (tRNA-guanine transglycosylase)	1.532
	121875	AA426299	Hs.98510	ESTs	1.532
30	114366	Z41747	Hs.469	succinate dehydrogenase complex; subunit A; flavoprotein (Fp)	1.531
	132944	AA054515	Hs.6127	ESTs; Weakly similar to prostate-specific transglutaminase [H.sapiens]	1.53
	111199	N68210	Hs.29822	ESTs	1.53
	113494	T88878	Hs.258738	ESTs	1.529
	129515	AA490882	Hs.112227	ESTs	1.528
35	133124	AA156049	Hs.65490	ESTs	1.528
	104785	AA027163	Hs.7942	ESTs	1.526
	105595	AA279408	Hs.25866	ESTs	1.526
	130198	U67156	Hs.151988	mitogen-activated protein kinase kinase kinase 5	1.526
40	114297	Z40758	Hs.173091	DKFZP434K151 protein	1.525
	112876	T03488	Hs.4842	ESTs	1.525
	127500	AA525014	Hs.162115	ESTs	1.525
	120519	AA258585	Hs.129887	cadherin 19 (NOTE: redefinition of symbol)	1.525
	119859	W80702	Hs.58461	ESTs	1.525
45	129944	L00389	Hs.1361	cytochrome P450; subfamily I (aromatic compound-inducible); polypeptide 2	1.524
	118864	N89670	Hs.42148	ESTs; Weakly similar to Su(P) [D.melanogaster]	1.523
	123964	C13961	Hs.210115	EST	1.523
	111676	R19414	Hs.166459	ESTs	1.522
	128332	AI079523	Hs.134173	ESTs	1.522
50	130455	X17059	Hs.155956	N-acetyltransferase 1 (arylamine N-acetyltransferase)	1.521
	125181	W58461	Hs.12396	ESTs	1.521
	127093	AA768241		oa72d02.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1317795 3', mRNA sequence.	1.521
	132156	AA157401	Hs.4113	S-adenosylhomocysteine hydrolase-like 1	1.521
55	125303	Z39821	Hs.107295	ESTs	1.52
	132697	AA281951	Hs.5518	Homo sapiens mRNA; cDNA DKFZp566J2146 (from clone DKFZp566J2146)	1.52
	117086	H93135	Hs.41840	ESTs	1.519
	113355	T79203	Hs.14480	ESTs	1.518
	108621	AA101811	Hs.69506	ESTs	1.518
60	109384	AA219172	Hs.86849	EST	1.518
	128510	X94703	Hs.100816	RAB28; member RAS oncogene family	1.517
	132968	N77151	Hs.61638	myosin X	1.515
	117035	H88798	Hs.41182	ESTs	1.515
	116781	H22985	Hs.52132	ESTs	1.513
	108677	AA115629	Hs.118531	ESTs	1.513
65	130214	H78003	Hs.15266	ESTs	1.513
	134700	AA481414	Hs.8868	golgi SNAP receptor complex member 1	1.512
	116618	D80783	Hs.45224	ESTs	1.508
	126257	N99638		tumor necrosis factor receptor superfamily; member 10b	1.508
	125859	AA806808	Hs.118797	ubiquitin-conjugating enzyme E2D 3 (homologous to yeast UBC4/5)	1.508

	113837	W57698	Hs.8888	ESTs	1.507
	114317	Z41038	Hs.469	succinate dehydrogenase complex; subunit A; flavoprotein (Fp)	1.507
	100311	D50640	Hs.184653	phosphodiesterase 3B; cGMP-inhibited	1.507
5	126802	AA947601	Hs.97056	ESTs	1.506
	128661	R82837	Hs.103329	KIAA0970 protein	1.506
	134194	AA233231	Hs.79828	ESTs	1.506
	108953	AA149652	Hs.42128	ESTs	1.504
	133240	D31161	Hs.68613	ESTs	1.502
10	132671	X76302	Hs.54649	putative nucleic acid binding protein RY-1	1.501
	132609	Z48923	Hs.53250	bone morphogenetic protein receptor; type II (serine/threonine kinase)	1.501
	105574	AA278678	Hs.258567	ESTs	1.5
	113718	T97782	Hs.256268	ESTs	1.5
	127824	AI208365	Hs.127811	ESTs	1.5
15	130132	U55936	Hs.184376	synaptosomal-associated protein; 23kD	1.5
	127394	AA453224		ESTs; Weakly similar to !! ALU SUBFAMILY J WARNING ENTRY !! [H.sapiens]	1.5
	100485	HG11111-HT1111		Ras-Like Protein Tc21	1.5
	101078	L04510	Hs.792	ADP-ribosylation factor domain protein 1; 64kD	1.5
	128611	AA456845	Hs.102471	KIAA0680 gene product	1.5

TABLE 12A shows the accession numbers for those primekeys lacking unigeneID's for Table 12. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Key:	Uniqua Eos probeset identifier number
CAT number:	Gene cluster number
Accession:	Genbank accession numbers

Pkey	CAT number	Accession
108536	119811_1	AA084524 AA339253 AW966289
117040	46956_1	AW970600 AA503323 H89218 AF086031 H89112
100782	18457_1	AA355435 NM_001516 Z30093 T28405 AW949486 AA461142 AA410532 AI652073 AA521208 AI970141 AI968234 AI026102 AA713583 AW135876 AA936614 AA770300 AI242635 AA377033 AW960263 AW607683 AI273603 AA410287 AI040513 AA460838 AI803916 AW294095 AW449680 AW798677 AW675048 BE542116 AL120521
100819	3022_1	L34840 NM_003241 U31905 AI546931 AI791616 AI973065 AI792321 AI546937 AI685880 AI732835 AI682360 AA420653 AA564047 AI682323 AI824614 AI659889 AI680052 AI970887 AI623108 AA420692 AI418074 AA631018 AI810595 AW291463 AW449930 AI668908 AI970818
100824	5_36	AI393237 AI521317 AI761348 AF025841 D43968 AW994987 L34598 AF025841 D89789 D89788 D89790 AW998932 AI971742 AI310238 X90976 AW139668 AW674280 AI365552 AA877452 AV657554 C75229 AA376077 AI798056 AW609213 W25586 H30149 BE075089 BE075190 AW580858 H99598 AA425238 AA133916 AW363478 BE158121 BE158127 AW467960 BE158135 BE158126 BE158145 N92860 AA847246 AI961688 AI361423 AA878154 AA043767 AI863712 AI559226 AW339007 AI371266 AI368901 AA046624 AA134739 AW449154 AA130232 AI458720 AA962511 AI700627 R70437 AW004008 AA045229 AI671572 H99599 AA043768 AI685454 AI871685 N29937 X90977 AA524240 AI142114 AI825750 AI567805 AI631365 AI347893 AA134740 F20669 AA046707 AW793216 AW963298 AW959380 AA363265 AI784593 AI268201 R69451 AV657618 AI695588 BE312163 AJ230798 AA374482 AI926059 AA622653 AI860704 BE139185 AW296884 T60238 T60120 U33921 AI190489 AA573311
125004	264197_1	AI814663 AA806761 AA765241 AA019317 AA092255 AA035405 T85079 AA890151 AI373959 T85080 BE153728 AA740848 BE080682 AL048137 AW182316 AI699468 AW274481 AW407538 AA306562 AW950024 AW949943 AL045703 AW843196 W25132 BE612794 AA304266 AW958054 H25673 AV646563 AV646573 BE172990 AW593488 AA385181 AA164998 AI246476 AA345406 AI277554 AA134749 AA856624 BE613247 AA299003 AL048138 AA028121 T92510 AI923835 AW020440 AI401594 AI889401 N93290 AA044247 AA028100 AI582845 AA811151 AI741811 AI925878 AA448277 AA172221 AI214783 BE220793 AA022746 AI082882 AA022849 AI928385 AA573472 AI420686 AW072902 AI799493 AI873506 AI468977 AI192079 AI468976 AA044272 AW015701 AW316979 AA933042 AA609017 AI318393 AI424571 AI934945 AA172023 AW050917 AA846180 AA134748 AI003947 AI766769 AW006697 AA653517 AW575680 AI474214 AA401478 U36922 AA927064 AA868000 D62654 T91745 AW500202 AA194764 AA746346 AA130464 AW117498 AA054526 N26432 H02534 H04964 AW303367 BE300931 AI218049 AI208073 AW182749 AA983630 AI147585 AA194765 AA054534 AA922720 AI436585 AI346535 AA134269 AA280923 AA897422 AA019559 AW274010 AA035406 AA917879 H99327 W32908 AI216046 AW496823 AA019414 H82288 W35284 AI936621 AI767113 AA866177 AW367874 H82398 AF032885 AW300151 AW467069 AA809346 AI188507 AI494178 AA872752 AI631631 U02310 NM_002015 AA815006 AI382453 AW197658 AI761654 AI804396 AI382221 AI813640 AI439635 AI523901 AW517242 AI221705 AW298104 AW204560 AW573095 AW028783 AW014650 AI766744 AI808294 AI698758 AI041809 AI766667 AI479103 AA872797 AA769305 AA765080 AA334166 AI472322
124704	292319_1	R07335 R07640
116988	185904_1	AW953679 AW953680 AA244436 H82527 AA361046 AA244483 H82526
124825	330773_1	AA501669 R52088
110455	46874_1	H52576 AF085971 H52172
126257	182217_1	N99638 AW973750 AA328271 H90994 AA558020 AA234435 N59599 R94815
125624	154135_1	AW968363 AA465492 R34539 AA165411
104038	264235_1	AA374532 AA421255
103427	43892_1	BE514383 AA071273 AW247987 AW673286 BE312102 AW749824 BE071985 AW577383 BE071945 BE072005 AW577355 BE071965 AW239231 BE072000 BE071960 AW577360 AW749830 AW373020 X97303 AW999522 BE000192 BE562219 BE266655 BE264970
104142	113242_1	AA074713 AA447006
127093	47721_1	AW977549 AA256038 AL365415 AW500455 AA768241 AW968097 Z17849 AA256104

	125873	10492_1	AW271838 AL133605 C01646 H29959 AA999896 D60676 AW999454 AW961176 AA315244 H14437 AW386118 N46512 AW272021 AI768516 BE466421 AI082809 AI804454 AA905101 AW173368 N38942 AW614169 AI080483 N29489 AI500550 AA994475 AA614464 AA707368 AA593145 AA569473 AW627815 AI828244 N63226 N42300
5	125954	4457_1	NM_016353 AB023584 W44753 R09585 AA382865 R23772 AI814257 AA974046 AK001608 AI935638 AW440609 AI420022 AA777386 AA806969 AI554876 AI584006 AI688556 AI688634 AI697997 AI014540 AI806683 AI741202 AW263154 AW297238 AI149951 AI589076 AW082158 AW614265 AA931887 AA781969 R09490 AA484643 AI207121 AI088390 AI538065 AI619547 AI741925 AI702846 H40846 R93943 AW747979 AA461348 U30163 AA326023 AI535992 AW242870 AI244025 AI222558 W38425 AW473630 AI624599 AI921226 AI683152 AI096458 AI123822 AW170802 C16447 AI337674 D25726 AW339366 AW771259 AA461174
10	125992	1589048_1	H48372 W01626
	127210	15307_6	AA305278 AA223833
15			110924 6443_1 AW058463 AF195766 AA680145 T86901 W60373 W60281 NM_007222 AF106862 AI000795 AA167188 AW884503 AW891313 AW891332 AW891312 AI984924 AI123518 N75170 AA131614 H25330 AI913358 AI742277 W25576 R58771 AW445159 AW888628 AW888627 AW274674 AI088482 N52314 N34282 AW001769 AI338943 T66784 AI288963 AW468676 AW237528 H25289 N71690 AA610128 AI143458 AI082599 N49144 AA854773 AW663411 AW610151 N47938 AA601626 AA167189 AA918304 AA805205 BE069496 AA652836 BE069499 AI699298 AW249926 AW888578 BE567635 T10726 AW604715 D54245 D53062 D55610 D55555 AA301376 AI133498 N77788 AI936320 AW090734 AI269977 N50828 AA550814 AI421993 AI005384 N50813 D60292 D59349 AA131710 D81698 D81699
20	127263	232161_1	AA331156 AA331157 AA331155
	135197	29440_1	U76456 NM_003256 AF057532 AA193414 AW293304 AW963378 AA313095 AI359841 AI969312 AI080163 AW448926 AI671136 BE466399 AI637967 AI671873 AW196583 AW071635 AI634427 AW296872 AW292470 AA193650
	127394	304844_1	BE161832 AA453224 AA485772
	126879	1860_2	D90391 M55575 AI652268 AA719776
25	126983	171841_1	AA524886 AW971347 AA211537
	120470	188975_1	AW971327 AA524988 AW628653 AA251797
	127854	443883_1	AW976796 AA769520
	121367	280429_1	AA432071 AA405648 AW000908 T16347
	106320	6435_1	AB028957 AL120001 AI267678 H10928 R19844 AW970334 AA393182 F05472 F11711 H09908 N50250 AI815411 BE463679 D61468 AW970253 D08889 C15548 D61011 D60867 AI815795 AA534831 D81386 AW235039 AI382158 D81174 AA416899 AA852310 H09789 H10929 H09813 F09369 R44721 D51515 Z38456 R14004 T66255 F12148 F12139 AW351702 M85350 AI018713 AW972450 AW972645 AA514964 T66172 F09785 F09776 AA436608 T05327 T07118 AA339352 AW301608 N46706 AA649093 AA287595 AW811753 AA287596 N39260 NM_001874 J04970 T91426 AW205201 T84979 AA255727 AA847837 R02164 T91339 AV651884 AV651835 AV651350 AV650118 AV651338 AI272002 AI367796 AA830651 AA262112 AW151198
30	115479	201515_1	AU076696 AA219720 AL135197 AA305877 N56376 AA318063 AA130725 AW954903 BE541230 AW383312 U86753 D85423
	101026	11075_1	AI679458 AI122932 AB007892 AI583919 BE160134 F08104 R34903 F13440 AA095444 AA262453 AA191036 R17895 T81266 BE149776 AI279537 AI143113 AA361072 AW959030 AW268817 AA811533 BE275179 AI221677 T65147 R49293 AA249176 BE000290 AA768053 F09494 BE092645 BE172099 Z41177 AA044750 AI909768 BE140795 BE140574 AW845210 AW752452 BE243244 AA843664 AI300080 BE169032 AW189979 BE004869 AA621872 AI951772 AI678897 AI926598 N62813 AI350912 AW608791 AI309602 AI983138 AW875592 AI655073 AW875626 AA130606 AI370827 C75528 C75554 AW263335 AI344426 BE004788 AA576220 AA604824 AI431405 AA749378 R38882 AW955075 AA173821 C75657 AA219672 AW768408 R43141 AI431414 AA483343 AI673792 T17294 AW770187 N74285 AI476404 AI088288 AA654152 AW974864 BE617311 BE243328 BE168049
35	100401	24827_1	U64675 AW167507 AW167508 BE218568 AA779360 W85722 AL044843 BE159404 AF012086 AW898611 AW898610 BE159405 BE092191 AW890826 AW369841 AW368064 AW606702 AL044731 R82691 AA419346 AA16558 H96045 AL040450 AI640531 AI808434 AL046613 AW855784 AW362469 AL048881 AL049015 AA094272 AA888908 AA417294 AW237786 R59793 AL044916 D82402 AI216854 AI079342 H96406 AL037845 AI915900 AA972133 AI478783 T31074 Z21135 Z21396 AA352182 R13918 AA430178 C17811 AI371824 AI742256 AA926801 N79156 AA350610 AA081971 N83639 R35544 AA312292 AW952080 N42322 AA171957 AA565297 R89207 AA504106 AI630782 AA826482 AI301579 T36241 AW966618 Z28426 AL043480 AI124636 AA393449 T19504 AW887823 AI289814 N53979 AL043571 AI632764 AI859613 AI986308 AI683212 AI984499 AI133258 C05898 AW512761 AI041260 BE466240 Z19161 AI351190 N67549 AI373374 AA400873 AW440914 AW514879 AA770146 AI358754 R51113 AI283773 AA649886 T30543 D54358 R37750 T03358 T15451 T15880 AA999689 N67396 AI056289 T85597 N62441 R89099 R00035 T85596 R61335 R00128 N63359 AI535964 AI207768 M31468 NM_012250 W01322 AA253280 AA253233 AA293148 AW582106 R79880 AA459547 AA363459 AA234396 N31669 H44468 AA434587 AW363088 AW993541
45	130542	28089_3	AA070906 AA070934
	108345	112277_6	X51501 NM_002652 Y10179 J03460 AI791618 AI821473 AA916588 AA564296 AA916110 AI972286 AI420470 AI568790
	100522	19669_1	AI597724 AW205207 AI659305 AI791620 AA532383 AI821475 AA526498
50	100533	32905_1	NM_012249 M31470 AL043108 AA262561 AA178883 T29433 AA313329 W48807 AW404323 AA453560 AW403227 H94816 W17101 AA165152 W23989 AA091310
	100598	23902_2	AL121734 D54896 AA424269 BE242906 AA362118 BE018454 AI280348 AL048769 M35543 AA757734 AI128865 H20289 H23728 AI203445 H41481 H18237 H44081 H92839 AI928621 H75675 D51148 AI796198 AW390453 D55579 D54145 D53996 D54015 R37664 H17541 AA668681 T65061 R15867 AW468123 R16049 H69030 AA054226 H16070 F09655 R92144 T03521 R05473 H92840 AA018186 R91707 U35637 AA112989 Z19308
65	102332	14745_3	
	118250	genbank_N62602	N62602
	103678	entrez_Z84483	Z84483
	119400	genbank_T92767	T92767
	119559	entrez_W38197	W38197

MISSING AT THE TIME OF PUBLICATION

TABLE 13: shows genes, including expression sequence tags, up-regulated in prostate tumor tissue compared to normal prostate tissue as analyzed using Affymetrix/Eos Hu02 GeneChip array. Shown are the ratios of "average" normal prostate to "average" prostate cancer tissues.

5

10 Pkey: Unique Eos probeset identifier number
 ExAccn: Exemplar Accession number, Genbank accession number
 UnigenelD: Unigene number
 Unigene Title: Unigene gene title
 R1: Background subtracted normal prostate : prostate tumor tissue

	Pkey	ExAccn	UnigenelD	Unigene Title	R1
15	333516			CH22_FGENES.173_1	0.028
	337954			CH22_EM:AC005500.GENSCAN.96-3	0.029
	332496	R73299	Hs.204354	ras homolog gene family; member B	0.03
	337944			CH22_EM:AC005500.GENSCAN.89-7	0.033
	334111			CH22_FGENES.330_10	0.033
20	333657			CH22_FGENES.241_2	0.034
	327718			CH.04_hs gil6525284	0.034
	336355			CH22_FGENES.817_5	0.035
	322011	AL137354		EST cluster (not in UniGene)	0.035
	336377			CH22_FGENES.821_5	0.036
25	300254	AW079607	Hs.188417	ESTs; Weakly similar to ZnT-3 [H.sapiens]	0.037
	330096			CH.19_p2 gil6015278	0.037
	335191			CH22_FGENES.507_6	0.038
	334040			CH22_FGENES.322_8	0.039
	333586			CH22_FGENES.204_2	0.04
30	333295			CH22_FGENES.132_2	0.042
	313326	AI088120	Hs.122329	ESTs	0.043
	329517			CH.10_p2 gil3983513	0.043
	333403			CH22_FGENES.144_21	0.043
	335226			CH22_FGENES.513_11	0.044
35	335976			CH22_FGENES.652_11	0.045
	333637			CH22_FGENES.229_2	0.046
	334582			CH22_FGENES.407_5	0.046
	336437			CH22_FGENES.826_4	0.047
	337461			CH22_FGENES.782-1	0.047
40	302892	N58545	Hs.6975	histone deacetylase 3	0.049
	338689			CH22_EM:AC005500.GENSCAN.475-3	0.049
	334721			CH22_FGENES.421_32	0.049
	305867	AA864572		EST singleton (not in UniGene) with exon hit	0.049
	335498			CH22_FGENES.571_7	0.05
45	311596	AI682088	Hs.223368	ESTs	0.05
	326959			CH.21_hs gil6469836	0.051
	311688	AW025661	Hs.240090	ESTs	0.052
	317298	AI922374	Hs.158549	ESTs	0.052
	332984			CH22_FGENES.54_6	0.052
50	321039	AW247083		EST cluster (not in UniGene)	0.053
	335844			CH22_FGENES.623_4	0.053
	325371			CH.12_hs gil5866920	0.054
	335667			CH22_FGENES.590_18	0.054
	333635			CH22_FGENES.228_2	0.054
55	336736			CH22_FGENES.110-2	0.055
	335893			CH22_FGENES.635_1	0.055
	333170			CH22_FGENES.94_5	0.055
	329768			CH.14_p2 gil6015501	0.055
	334030			CH22_FGENES.320_2	0.055
60	323359	AA234172	Hs.137418	ESTs	0.055
	300453	AW051431	Hs.113029	ribosomal protein S25	0.055
	334262			CH22_FGENES.367_12	0.055
	306590	AI000246		EST singleton (not in UniGene) with exon hit	0.055
	331087	R22520	Hs.23398	ESTs	0.055
65	338620			CH22_EM:AC005500.GENSCAN.450-18	0.056
	339045			CH22_DA59H18.GENSCAN.28-5	0.056
	308023	AI452732		EST singleton (not in UniGene) with exon hit	0.057

	339067		CH22_DA59H18.GENSCAN.33-3	0.057
	335689		CH22_FGENES.596_4	0.057
	339069		CH22_DA59H18.GENSCAN.33-5	0.057
5	338176		CH22_EM:AC005500.GENSCAN.219-4	0.057
	328159		CH.06_hs gj 5868065	0.058
	335655		CH22_FGENES.590_6	0.058
	336371		CH22_FGENES.820_1	0.058
	336558		CH22_FGENES.842_3	0.059
10	337738		CH22_EM:AC000097.GENSCAN.100-4	0.059
	334273		CH22_FGENES.369_2	0.059
	335889		CH22_FGENES.633_3	0.059
	327807		CH.05_hs gj 5867968	0.059
	333315		CH22_FGENES.138_7	0.059
15	338825		CH22_DJ246D7.GENSCAN.4-6	0.06
	337612		CH22_C20H12.GENSCAN.22-5	0.06
	333897		CH22_FGENES.293_4	0.06
	335990		CH22_FGENES.655_4	0.06
	334264		CH22_FGENES.367_15	0.06
20	338653		CH22_EM:AC005500.GENSCAN.460-39	0.061
	322303	W07459	EST cluster (not in UniGene)	0.061
	333498		CH22_FGENES.168_8	0.061
	336522		CH22_FGENES.839_3	0.061
	301357	AW295677	Hs.137840 ESTs; Moderately similar to HOMEOBOX PROTEIN SIX1 [H.sapiens]	0.062
25	305917	AA876469	Hs.181357 laminin receptor 1 (67kD; ribosomal protein SA)	0.062
	336143		CH22_FGENES.705_5	0.063
	333493		CH22_FGENES.168_2	0.063
	332533	M99487	Hs.1915 folate hydrolase (prostate-specific membrane antigen) 1	0.063
30	325844		CH.16_hs gj 6552453	0.063
	336402		CH22_FGENES.823_17	0.063
	335767		CH22_FGENES.607_1	0.064
	301893	T80334	EST cluster (not in UniGene) with exon hit	0.064
	324019	AW177009	EST cluster (not in UniGene)	0.064
35	305801	AA845997	EST singleton (not in UniGene) with exon hit	0.064
	335188		CH22_FGENES.507_3	0.065
	337533		CH22_FGENES.828-2	0.065
	333311		CH22_FGENES.138_3	0.065
	335668		CH22_FGENES.590_19	0.065
40	306786	AI041589	EST singleton (not in UniGene) with exon hit	0.066
	306365	AA962086	EST singleton (not in UniGene) with exon hit	0.066
	306249	AA933840	EST singleton (not in UniGene) with exon hit	0.066
	335018		CH22_FGENES.474_6	0.066
	333594		CH22_FGENES.210_3	0.066
45	333900		CH22_FGENES.293_7	0.066
	325207		CH.10_hs gj 6552430	0.067
	329888		CH.15_p2 gj 6067149	0.067
	326238		CH.17_hs gj 5867260	0.067
	333658		CH22_FGENES.241_4	0.067
50	335809		CH22_FGENES.617_6	0.068
	307427	AI243437	EST singleton (not in UniGene) with exon hit	0.068
	318428	AI949409	Hs.224583 ESTs	0.069
	327005		CH.21_hs gj 5867664	0.069
	330463	HG998-HT998	Sulfotransferase, Phenol-Preferring	0.069
55	333318		CH22_FGENES.138_10	0.07
	333313		CH22_FGENES.138_5	0.07
	325937		CH.16_hs gj 5867132	0.07
	335663		CH22_FGENES.590_14	0.07
	335349		CH22_FGENES.539_2	0.07
60	303396	AA224470	Hs.25426 ESTs; Weakly similar to unknown [H.sapiens]	0.07
	332603	N66681	Hs.33470 ESTs	0.07
	333310		CH22_FGENES.138_2	0.071
	309924	AW340812	EST singleton (not in UniGene) with exon hit	0.071
	336340		CH22_FGENES.814_15	0.071
65	308025	AI453365	Hs.172928 collagen; type I; alpha 1	0.071
	306805	AI055966	EST singleton (not in UniGene) with exon hit	0.071
	335499		CH22_FGENES.571_8	0.071
	329669		CH.14_p2 gj 6272129	0.071
	321666	D28390	EST cluster (not in UniGene)	0.071
	338174		CH22_EM:AC005500.GENSCAN.219-2	0.072

	336556		CH22_FGENES.842_1	0.072
	305451	AA738105	Hs.140 immunoglobulin gamma 3 (Gm marker)	0.072
	336684		CH22_FGENES.46-1	0.072
5	326943		CH.21_hs gij6004446	0.073
	333947		CH22_FGENES.303_1	0.074
	333214		CH22_FGENES.104_5	0.074
	331917	AA446572	Hs.174007 ESTs; Moderately similar to !!!! ALU SUBFAMILY J WARNING	0.074
	339102		CH22_DA59H18.GENSCAN.44-9	0.074
10	328122		CH.06_hs gij5868031	0.075
	332250	N62712	Hs.226223 KIAA0618 gene product	0.075
	328506		CH.07_hs gij5868471	0.075
	331756	AA291468	Hs.98504 ESTs	0.075
	335193		CH22_FGENES.507_8	0.076
15	317729	AA971718	Hs.128141 ESTs	0.076
	304515	AA458708	Hs.251577 hemoglobin; alpha 2	0.076
	313644	AI565766	Hs.124960 ESTs	0.076
	326145		CH.17_hs gij5867204	0.076
	336394		CH22_FGENES.823_6	0.077
20	306516	AA989542	EST singleton (not in UniGene) with exon hit	0.077
	300629	AA152119	Hs.155101 ATP synthase; H+ transporting; mitochondrial F1 complex; alpha subunit; isoform 1; cardiac muscle	0.077
	333160		CH22_FGENES.91_2	0.077
	337490		CH22_FGENES.799-5	0.077
25	305403	AA723748	EST singleton (not in UniGene) with exon hit	0.077
	331747	AA281765	Hs.193689 ESTs	0.077
	332792		CH22_FGENES.3_2	0.078
	330513	M81057	Hs.180884 carboxypeptidase B1 (tissue)	0.078
	308905	AI859636	Hs.8102 ribosomal protein S20	0.078
30	337419		CH22_FGENES.759-4	0.078
	333459		CH22_FGENES.157_8	0.078
	334851		CH22_FGENES.440_3	0.078
	329046		CH.X_hs gij5868569	0.078
	327879		CH.06_hs gij5868142	0.079
35	305830	AA857665	EST singleton (not in UniGene) with exon hit	0.079
	302928	AL137719	EST cluster (not in UniGene) with exon hit	0.079
	304321	AA136698	Hs.113029 ribosomal protein S25	0.079
	326390		CH.19_hs gij5867340	0.079
	335230		CH22_FGENES.514_2	0.08
40	334622		CH22_FGENES.412_6	0.08
	335331		CH22_FGENES.535_4	0.08
	304753	AA578840	Hs.77961 major histocompatibility complex; class I; B	0.08
	301863	AI418863	EST cluster (not in UniGene) with exon hit	0.081
	336561		CH22_FGENES.842_6	0.081
45	335611		CH22_FGENES.583_5	0.081
	305060	AA635771	EST singleton (not in UniGene) with exon hit	0.081
	306051	AA905130	EST singleton (not in UniGene) with exon hit	0.082
	308289	AI571211	EST singleton (not in UniGene) with exon hit	0.082
	334365		CH22_FGENES.378_13	0.082
50	335496		CH22_FGENES.571_4	0.082
	332634	S38953	Human unidentified gene complementary to P450c21 gene; partial cds	0.082
	337824		CH22_EM:AC005500.GENSCAN.13-18	0.082
	335822		CH22_FGENES.619_7	0.082
55	334758		CH22_FGENES.428_7	0.082
	309641	AW194230	Hs.253100 EST	0.082
	333064		CH22_FGENES.75_7	0.083
	338695		CH22_EM:AC005500.GENSCAN.477-25	0.083
	331809	AA402482	Hs.97312 ESTs	0.083
60	326138		CH.17_hs gij5867203	0.083
	328304		CH.07_hs gij6004478	0.083
	330570	U60276	Hs.165439 arsA (bacterial) arsenite transporter; ATP-binding; homolog 1	0.083
	334305		CH22_FGENES.373_8	0.083
	335885		CH22_FGENES.632_3	0.083
65	325839		CH.16_hs gij6552452	0.083
	333531		CH22_FGENES.175_18	0.084
	330385	AA449749	Hs.31386 ESTs; Highly similar to secreted apoptosis related protein 1 [H.sapiens]	0.084
	323305	AA811351	Hs.25307 Homo sapiens clone 24812 mRNA sequence	0.084
	331698	Z39929	Hs.65843 ESTs	0.084

	335888		CH22_FGENES.633_2	0.084
	306008	AA894390	EST singleton (not in UniGene) with exon hit	0.084
	334249		CH22_FGENES.365_15	0.084
5	318303	AW451197	Hs.113418 ESTs	0.084
	330171		CH.02_p2 gij6648220	0.084
	336662		CH22_FGENES.41-1	0.085
	320506	AI815668	Hs.157476 suc1-associated neurotrophic factor target 2 (FGFR signalling adaptor)	0.085
10	316974	AI740721	Hs.128292 ESTs	0.085
	336492		CH22_FGENES.832_9	0.085
	335750		CH22_FGENES.602_4	0.085
	335676		CH22_FGENES.594_1	0.086
	336093		CH22_FGENES.691_2	0.086
15	310932	AI933861	Hs.222852 ESTs	0.086
	335160		CH22_FGENES.502_4	0.086
	334306		CH22_FGENES.373_9	0.086
	334793		CH22_FGENES.433_5	0.086
	333936		CH22_FGENES.301_2	0.087
20	336413		CH22_FGENES.823_35	0.087
	333775		CH22_FGENES.272_6	0.087
	335971		CH22_FGENES.652_4	0.087
	301737	AI815981	EST cluster (not in UniGene) with exon hit	0.087
25	339101		CH22_DA59H18.GENSCAN.44-6	0.087
	327612		CH.04_hs gij6525283	0.087
	326241		CH.17_hs gij5867260	0.088
	338386		CH22_EM:AC005500.GENSCAN.331-4	0.088
	327762		CH.05_hs gij5867961	0.088
	305266	AA679772	EST singleton (not in UniGene) with exon hit	0.088
30	334359		CH22_FGENES.378_4	0.088
	335500		CH22_FGENES.571_10	0.088
	329687		CH.14_p2 gij6117856	0.088
	333654		CH22_FGENES.240_2	0.088
	324430	AA464018	EST cluster (not in UniGene)	0.088
35	325999		CH.16_hs gij5867073	0.089
	334832		CH22_FGENES.439_1	0.089
	339115		CH22_DA59H18.GENSCAN.49-3	0.089
	300896	AI916902	Hs.213882 ESTs	0.089
40	328784		CH.07_hs gij5868309	0.089
	335044		CH22_FGENES.480_1	0.089
	329791		CH.14_p2 gij6469354	0.089
	333656		CH22_FGENES.240_4	0.089
	326180		CH.17_hs gij5867211	0.089
	333391		CH22_FGENES.144_6	0.089
45	338324		CH22_EM:AC005500.GENSCAN.306-3	0.089
	305396	AA721052	EST singleton (not in UniGene) with exon hit	0.089
	337483		CH22_FGENES.795-7	0.09
	326424		CH.19_hs gij5867369	0.09
	306454	AA977992	EST singleton (not in UniGene) with exon hit	0.09
50	338893		CH22_DJ32H10.GENSCAN.7-6	0.09
	327470		CH.02_hs gij5867772	0.09
	333165		CH22_FGENES.91_7	0.09
	307155	AI186738	Hs.182426 ribosomal protein S2	0.09
	330717	AA233926	Hs.23635 ESTs	0.09
55	335334		CH22_FGENES.535_10	0.09
	335907		CH22_FGENES.636_2	0.09
	333885		CH22_FGENES.292_7	0.09
	331034	N51868	Hs.31965 ESTs; Moderately similar to 40S RIBOSOMAL PROTEIN S20 [H.sapiens]	0.09
60	304660	AA534416	Hs.162185 ESTs	0.09
	328217		CH.06_hs gij5868096	0.091
	336068		CH22_FGENES.684_13	0.091
	302833	AA295381	Hs.44423 ESTs	0.091
	328668		CH.07_hs gij5868254	0.091
65	335309		CH22_FGENES.532_2	0.091
	338481		CH22_EM:AC005500.GENSCAN.377-5	0.091
	306286	AA936892	EST singleton (not in UniGene) with exon hit	0.091
	305070	AA639783	EST singleton (not in UniGene) with exon hit	0.091
	304870	AA594811	Hs.119122 ribosomal protein L13a	0.091
	303856	AA968589	Hs.944 glucose phosphate isomerase	0.091

5	323789	AI459812	Hs.170460	ESTs; Weakly similar to KIAA0990 protein [H.sapiens]	0.092
	334910			CH22_FGENES.455_3	0.092
	326382			CH.19_hs gil5867327	0.092
	332467	AA489630	Hs.119004	KIAA0665 gene product	0.092
	338534			CH22_EM:AC005500.GENSCAN.402-7	0.092
10	336449			CH22_FGENES.829_6	0.092
	333709			CH22_FGENES.250_24	0.092
	336559			CH22_FGENES.842_4	0.092
	333230			CH22_FGENES.107_10	0.093
	333133			CH22_FGENES.83_9	0.093
15	334885			CH22_FGENES.451_11	0.093
	330605	X02419	Hs.77274	plasminogen activator; urokinase	0.093
	336392			CH22_FGENES.823_4	0.093
	334083			CH22_FGENES.327_38	0.093
	325469			CH.12_hs gil6017034	0.093
20	331077	R09531	Hs.19039	ESTs	0.093
	303701	AW500732		EST cluster (not in UniGene) with exon hit	0.093
	334218			CH22_FGENES.358_3	0.093
	336542			CH22_FGENES.840_6	0.093
	337151			CH22_FGENES.546-1	0.093
25	333642			CH22_FGENES.231_2	0.093
	336863			CH22_FGENES.297-4	0.093
	334680			CH22_FGENES.419_2	0.093
	326365			CH.18_hs gil5867297	0.093
	338952			CH22_DJ32110.GENSCAN.23-22	0.093
30	337539			CH22_FGENES.832-4	0.094
	333546			CH22_FGENES.180_2	0.094
	335258			CH22_FGENES.518_3	0.094
	336786			CH22_FGENES.168-19	0.094
	321644	AI204177	Hs.237396	ESTs	0.094
35	335943			CH22_FGENES.646_17	0.094
	327918			CH.06_hs gil5868165	0.094
	306398	AA970548		EST singleton (not in UniGene) with exon hit	0.094
	335671			CH22_FGENES.592_3	0.094
	335033			CH22_FGENES.475_11	0.094
40	338277			CH22_EM:AC005500.GENSCAN.290-2	0.094
	332061	AA504812	Hs.192824	early B-cell factor	0.094
	305153	AA654582	Hs.77039	ribosomal protein S3A	0.094
	333880			CH22_FGENES.292_2	0.094
	323940	AI864428	Hs.170880	ESTs	0.094
45	313779	AA648796	Hs.129771	ESTs	0.095
	323109	AA169345		EST cluster (not in UniGene)	0.095
	332930			CH22_FGENES.38_4	0.095
	335368			CH22_FGENES.543_6	0.095
	303887	R72672	Hs.193484	ESTs; Weakly similar to Similarity with yeast gene	0.095
50	336223			L3502.1 [C.elegans]	0.095
	311280	AI767957	Hs.197737	ESTs; Weakly similar to Y38A8.1 gene product [C.elegans]	0.095
	337256			CH22_FGENES.648-3	0.095
	308814	AI819263		EST singleton (not in UniGene) with exon hit	0.095
	334659			CH22_FGENES.418_7	0.095
55	335895			CH22_FGENES.635_3	0.095
	321697	AW388061	Hs.4953	golgi autoantigen; golgin subfamily a; 3	0.095
	336010			CH22_FGENES.668_8	0.096
	302824	U21260		EST cluster (not in UniGene) with exon hit	0.096
	333612			CH22_FGENES.217_7	0.096
60	304823	AA584837		EST singleton (not in UniGene) with exon hit	0.096
	335665			CH22_FGENES.590_16	0.096
	306518	AA989598		EST singleton (not in UniGene) with exon hit	0.096
	335243			CH22_FGENES.516_4	0.096
	335436			CH22_FGENES.559_5	0.096
65	300243	AI420256	Hs.161271	ESTs	0.096
	332810			CH22_FGENES.7_12	0.097
	308612	AI735634		EST singleton (not in UniGene) with exon hit	0.097
	335818			CH22_FGENES.618_6	0.097
	325838			CH.16_hs gil6552452	0.097
	337482			CH22_FGENES.795-6	0.097
	336645			CH22_FGENES.26-1	0.097
	337293			CH22_FGENES.675-1	0.098

	329893		CH.15_p2 gij6525313	0.098
	326533		CH.19_hs gij5867441	0.098
	334905		CH22_FGENES.452_20	0.098
5	306347	AA961144	EST singleton (not in UniGene) with exon hit	0.098
	336676		CH22_FGENES.43-4	0.098
	339166		CH22_DA59H18.GENSCAN.69-7	0.098
	335774		CH22_FGENES.607_10	0.098
	339216		CH22_FF113D11.GENSCAN.6-11	0.098
10	335311		CH22_FGENES.532_4	0.098
	329632		CH.11_p2 gij6729060	0.098
	328595		CH.07_hs gij5868224	0.098
	326928		CH.21_hs gij6456782	0.098
	315234	AI079680	Hs.120770 ESTs	0.098
	306082	AA908508	EST singleton (not in UniGene) with exon hit	0.098
15	305710	AA826544	EST singleton (not in UniGene) with exon hit	0.098
	318540	T30280	EST cluster (not in UniGene)	0.099
	337553		CH22_C4G1.GENSCAN.2-1	0.099
	320951	AA344069	Hs.202699 neurexophilin 4	0.099
	303845	T08033	EST cluster (not in UniGene) with exon hit	0.099
20	338981		CH22_DA59H18.GENSCAN.2-5	0.099
	321313	R87365	Hs.26058 ESTs; Weakly similar to p53 [H.sapiens]	0.099
	328348		CH.07_hs gij5868383	0.099
	332203	H49388	Hs.102082 EST	0.099
	301780	R07064	EST cluster (not in UniGene) with exon hit	0.099
25	332095	AA608838	Hs.162681 EST	0.099
	333227		CH22_FGENES.107_5	0.099
	316442	AA760894	Hs.153023 ESTs	0.099
	326001		CH.16_hs gij5867073	0.099
30	334363		CH22_FGENES.378_11	0.099
	338895		CH22_DJ32110.GENSCAN.9-2	0.099
	327460		CH.02_hs gij6004455	0.099
	332705	T59161	Hs.76293 thymosin; beta 10	0.1
	307806	AI351739	EST singleton (not in UniGene) with exon hit	0.1
35	322800	F25037	Hs.225175 ESTs	0.1
	304918	AA602697	EST singleton (not in UniGene) with exon hit	0.1
	334327		CH22_FGENES.375_4	0.1
	318359	AI097439	Hs.135548 ESTs	0.1
	326644		CH.20_hs gij5867559	0.1
40	334454		CH22_FGENES.388_3	0.1
	327959		CH.06_hs gij5868210	0.1
	323783	AA330586	Hs.131819 ESTs	0.1
	309198	AI955915	Hs.248038 major histocompatibility complex; class I; C	0.1
	339265		CH22_BA354112.GENSCAN.10-3	0.1
45	320576	AL049977	Hs.162209 Homo sapiens mRNA; cDNA DKFZp564C122 (from clone DKFZp564C122)	0.1
	338132		CH22_EM:AC005500.GENSCAN.200-2	0.1
	333163		CH22_FGENES.91_5	0.101
	337584		CH22_C20H12.GENSCAN.5-1	0.101
50	307588	AI285535	EST singleton (not in UniGene) with exon hit	0.101
	336969		CH22_FGENES.378-2	0.101
	327535		CH.02_hs gij6525279	0.101
	328732		CH.07_hs gij5868289	0.101
	336686		CH22_FGENES.46-3	0.101
55	335777		CH22_FGENES.607_13	0.101
	332944		CH22_FGENES.47_3	0.101
	333174		CH22_FGENES.95_1	0.101
	336380		CH22_FGENES.821_8	0.101
	330571	U60800	Hs.79089 sema domain; immunoglobulin domain (Ig); cytoplasmic domain; (semaphorin) 4D	0.101
60	331789	AA398721	Hs.186749 ESTs	0.101
	338915		CH22_DJ32110.GENSCAN.12-1	0.101
	334844		CH22_FGENES.439_24	0.101
	336642		CH22_FGENES.23-4	0.101
	334906		CH22_FGENES.452_21	0.101
65	333188		CH22_FGENES.98_8	0.101
	300088	AW299993	EST cluster (not in UniGene) with exon hit	0.101
	329373		CH.X_hs gij6682537	0.102
	331120	R46576	Hs.23239 ESTs	0.102
	335856		CH22_FGENES.628_1	0.102

	331888	AA431337	Hs.98017	ESTs	0.102
	333154			CH22_FGENES.89_4	0.102
	335989			CH22_FGENES.655_2	0.102
5	304385	AA235602		EST singleton (not in UniGene) with exon hit	0.102
	338016			CH22_EM:AC005500.GENSCAN.133-1	0.102
	335190			CH22_FGENES.507_5	0.102
	318595	T39486	Hs.6137	ESTs	0.102
	333697			CH22_FGENES.250_11	0.102
10	306526	AA989713		EST singleton (not in UniGene) with exon hit	0.103
	328734			CH.07_hs gij5868289	0.103
	307294	AI205612	Hs.73742	ribosomal protein; large; P0	0.103
	327424			CH.02_hs gij5867751	0.103
	335872			CH22_FGENES.630_3	0.103
15	333572			CH22_FGENES.189_1	0.103
	334774			CH22_FGENES.430_6	0.103
	338660			CH22_EM:AC005500.GENSCAN.462-1	0.103
	326713			CH.20_hs gij5867595	0.103
	333994			CH22_FGENES.310_18	0.103
	335800			CH22_FGENES.613_4	0.103
20	318113	AI187943	Hs.132322	ESTs	0.103
	337278			CH22_FGENES.665-1	0.103
	336386			CH22_FGENES.822_6	0.103
	334790			CH22_FGENES.432_15	0.103
25	303778	AW505368		EST cluster (not in UniGene) with exon hit	0.104
	336524			CH22_FGENES.839_5	0.104
	328936			CH.08_hs gij5868500	0.104
	335102			CH22_FGENES.494_7	0.104
	300935	AA513644	Hs.222815	ESTs; Weakly similar to Wiskott-Aldrich Syndrome protein [H.sapiens]	0.104
30	307581	AI284415		EST singleton (not in UniGene) with exon hit	0.104
	317301	AW291683	Hs.226056	ESTs	0.104
	335330			CH22_FGENES.535_3	0.104
	337968			CH22_EM:AC005500.GENSCAN.103-2	0.104
35	335627			CH22_FGENES.584_7	0.104
	336274			CH22_FGENES.762_2	0.104
	334730			CH22_FGENES.424_5	0.105
	334409			CH22_FGENES.383_6	0.105
	327237			CH.01_hs gij5867544	0.105
40	333321			CH22_FGENES.138_13	0.105
	303181	AA452366		EST cluster (not in UniGene) with exon hit	0.105
	333738			CH22_FGENES.261_2	0.105
	338255			CH22_EM:AC005500.GENSCAN.276-3	0.105
	334282			CH22_FGENES.369_12	0.105
	330190			CH.05_p2 gij6165182	0.105
45	310748	AW014249	Hs.158698	ESTs	0.105
	338150			CH22_EM:AC005500.GENSCAN.207-2	0.105
	336719			CH22_FGENES.82-6	0.105
	330228			CH.05_p2 gij6013527	0.105
50	327801			CH.05_hs gij5867924	0.105
	330525	S75168	Hs.274	megakaryocyte-associated tyrosine kinase	0.105
	334972			CH22_FGENES.468_2	0.105
	335111			CH22_FGENES.494_19	0.106
	334483			CH22_FGENES.395_5	0.106
55	328829			CH.07_hs gij5868337	0.106
	302753	M74299		EST cluster (not in UniGene) with exon hit	0.106
	334512			CH22_FGENES.398_10	0.106
	330024			CH.16_p2 gij6671908	0.106
	321030	AI769930	Hs.233617	Homo sapiens (clone B3B3E13) Huntington's disease candidate region	0.107
60	338410			CH22_EM:AC005500.GENSCAN.341-6	0.107
	334353			CH22_FGENES.376_5	0.107
	338276			CH22_EM:AC005500.GENSCAN.288-9	0.107
	329053			CH.X_hs gij5868574	0.107
	336560			CH22_FGENES.842_5	0.107
65	332158	AA621363	Hs.112980	EST	0.107
	336447			CH22_FGENES.829_4	0.107
	333703			CH22_FGENES.250_17	0.107
	326207			CH.17_hs gij5867222	0.107
	333232			CH22_FGENES.108_1	0.107

	334802		CH22_FGENES.435_1	0.107
	303784	AA704983	EST cluster (not in UniGene) with exon hit	0.107
	338847		CH22_DJ246D7.GENSCAN.10-2	0.107
5	339407		CH22_DJ579N16.GENSCAN.1-9	0.108
	337635		CH22_C20H12.GENSCAN.32-8	0.108
	334650		CH22_FGENES.417_17	0.108
	308511	AI687580	EST singleton (not in UniGene) with exon hit	0.108
	333392		CH22_FGENES.144_8	0.108
	325840		CH.16_hs gij6552452	0.108
10	315044	AW205664	Hs.129568 ESTs	0.108
	333298		CH22_FGENES.133_4	0.108
	335157		CH22_FGENES.501_7	0.108
	333305		CH22_FGENES.137_2	0.108
	326379		CH.19_hs gij5867327	0.108
15	335050		CH22_FGENES.482_1	0.108
	305185	AA663985	Hs.248038 major histocompatibility complex; class I; C	0.108
	335658		CH22_FGENES.590_9	0.108
	323040	AA336609	Hs.10862 ESTs	0.108
	337326		CH22_FGENES.699-6	0.108
20	339262		CH22_BA354112.GENSCAN.9-6	0.108
	321202	H54052	Hs.163639 ESTs; Weakly similar to INTERCELLULAR ADHESION MOLECULE-1 PRECURSOR [H.sapiens]	0.109
	331792	AA398968	Hs.97548 EST	0.109
	333806		CH22_FGENES.278_2	0.109
25	321325	AB033100	EST cluster (not in UniGene)	0.109
	331373	AA435513	Hs.178170 ESTs; Weakly similar to DUAL SPECIFICITY PROTEIN PHOSPHATASE 3	0.87
	328775		CH.07_hs gij5868309	0.109
	335105		CH22_FGENES.494_10	0.109
30	300975	AI283548	Hs.149668 ESTs	0.109
	324893	T31940	EST cluster (not in UniGene)	0.109
	333397		CH22_FGENES.144_15	0.109
	336484		CH22_FGENES.831_3	0.109
	335507		CH22_FGENES.571_22	0.109
35	336373		CH22_FGENES.820_3	0.109
	336188		CH22_FGENES.717_12	0.109
	313455	AW081702	Hs.137329 ESTs	0.109
	335185		CH22_FGENES.506_4	0.109
40	306814	AI066577	EST singleton (not in UniGene) with exon hit	0.109
	311130	AI632322	Hs.195306 ESTs	0.109
	310882	AW080339	Hs.211911 ESTs	0.109
	323383	AI346359	Hs.135209 ESTs	0.11
	300212	AW135925	Hs.184552 biphenylhydrolase-like (serine hydrolase; breast epithelial mucin-assoc.	0.11
45	325675		CH.14_hs gij5867014	0.11
	330095		CH.19_p2 gij6015278	0.11
	331942	AA453261	Hs.99309 ESTs	0.11
	334723		CH22_FGENES.421_34	0.11
	333614		CH22_FGENES.217_9	0.11
50	337316		CH22_FGENES.692-1	0.11
	305057	AA635626	Hs.62954 ferritin; heavy polypeptide 1	0.11
	338704		CH22_EM:AC005500.GENSCAN.480-3	0.11
	335385		CH22_FGENES.543_27	0.11
	338012		CH22_EM:AC005500.GENSCAN.128-10	0.11
55	329449		CH.Y_hs gij5868886	0.11
	338980		CH22_DA59H18.GENSCAN.2-4	0.11
	336553		CH22_FGENES.841_10	0.111
	330021		CH.16_p2 gij6671889	0.111
	327579		CH.03_hs gij5867824	0.111
60	333099		CH22_FGENES.79_4	0.111
	337076		CH22_FGENES.453-4	0.111
	331388	AA456852	Hs.43543 suppressor of white apricot homolog 2	0.111
	306674	AI005542	Hs.180414 heat shock 70kD protein 10 (HSC71)	0.111
	305949	AA884409	EST singleton (not in UniGene) with exon hit	0.111
65	330748	AA419217	Hs.15911 DKFZP586E1422 protein	0.111
	333780		CH22_FGENES.273_2	0.111
	323676	AI702835	EST cluster (not in UniGene)	0.111
	308952	AI868157	Hs.224226 EST	0.111
	309338	AW026946	Hs.181165 eukaryotic translation elongation factor 1 alpha 1	0.111

	329317		CH.X_hs gi 6381976	0.112
	333518		CH22_FGENES.173_3	0.112
	306982	AI127883	EST singleton (not in UniGene) with exon hit	0.112
5	336225		CH22_FGENES.728_2	0.112
	333698		CH22_FGENES.250_12	0.112
	302173	AI417947	Hs.14068 ESTs	0.112
	335510		CH22_FGENES.571_25	0.112
	328042		CH.06_hs gi 5902482	0.112
10	336512		CH22_FGENES.834_7	0.112
	328541		CH.07_hs gi 5868486	0.112
	311265	AW205118	Hs.199214 ESTs	0.112
	323218	AF131846	Hs.13396 Homo sapiens clone 25028 mRNA sequence	0.112
	302002	AF013956	Hs.123085 chromobox homolog 4 (Drosophila Pc class)	0.112
	315088	AA557351	Hs.152448 ESTs; Moderately similar to MULTIFUNCTIONAL PROTEIN ADE2	0.112
15	312581	AI937242	Hs.176590 ESTs	0.112
	322246	AW384710	Hs.125258 ESTs	0.112
	333659		CH22_FGENES.241_5	0.113
	327510		CH.02_hs gi 6117815	0.113
	336520		CH22_FGENES.839_1	0.113
20	338682		CH22_EM:AC005500.GENSCAN.472-1	0.113
	334508		CH22_FGENES.398_6	0.113
	322533	T59538	EST cluster (not in UniGene)	0.113
	306873	AI086929	EST singleton (not in UniGene) with exon hit	0.113
	336040		CH22_FGENES.679_2	0.113
25	303898	T23215	EST cluster (not in UniGene) with exon hit	0.113
	312011	AW294868	Hs.187226 ESTs	0.113
	335186		CH22_FGENES.506_5	0.113
	333607		CH22_FGENES.216_2	0.113
30	305549	AA773530	EST singleton (not in UniGene) with exon hit	0.113
	333686		CH22_FGENES.249_4	0.113
	334352		CH22_FGENES.376_3	0.113
	338195		CH22_EM:AC005500.GENSCAN.233-18	0.114
	333588		CH22_FGENES.206_2	0.114
35	339233		CH22_BA354112.GENSCAN.2-3	0.114
	337455		CH22_FGENES.777-1	0.114
	309101	AI925108	EST singleton (not in UniGene) with exon hit	0.114
	328522		CH.07_hs gi 5868477	0.114
	323999	AI537333	Hs.252782 ESTs	0.114
40	333517		CH22_FGENES.173_2	0.114
	329935		CH.16_p2 gi 6165200	0.114
	326226		CH.17_hs gi 5867230	0.114
	335890		CH22_FGENES.633_4	0.114
	336715		CH22_FGENES.77-1	0.114
45	327640		CH.04_hs gi 5867890	0.114
	338842		CH22_DJ246D7.GENSCAN.7-1	0.114
	306534	AA991487	EST singleton (not in UniGene) with exon hit	0.114
	336597		CH22_FGENES.266_1	0.114
	321010	Y17456	Hs.227150 Homo sapiens LSFR2 gene; last exon	0.114
50	302294	AA159213	Hs.5337 isocitrate dehydrogenase 2 (NADP+); mitochondrial	0.114
	324895	N44238	Hs.77515 inositol 1;4;5-triphosphate receptor; type 3	0.114
	327358		CH.01_hs gi 6552411	0.114
	308792	AI815153	Hs.195188 glyceraldehyde-3-phosphate dehydrogenase	0.115
	325886		CH.16_hs gi 5867087	0.115
	336850		CH22_FGENES.272-11	0.115
55	305858	AA863103	EST singleton (not in UniGene) with exon hit	0.115
	302569	AC004472	multiple UniGene matches	0.115
	336158		CH22_FGENES.707_2	0.115
	327866		CH.06_hs gi 5868131	0.115
60	339157		CH22_DA59H18.GENSCAN.67-3	0.115
	339258		CH22_BA354112.GENSCAN.8-3	0.115
	336129		CH22_FGENES.701_17	0.115
	333684		CH22_FGENES.249_2	0.115
	309618	AW190162	Hs.184776 ribosomal protein L23a	0.115
65	312926	AA954097	Hs.127523 ESTs	0.115
	302640	AB035698	EST cluster (not in UniGene) with exon hit	0.115
	328968		CH.08_hs gi 6456775	0.115
	327902		CH.06_hs gi 5868158	0.115
	321927	AJ223366	EST cluster (not in UniGene)	0.115
	335962		CH22_FGENES.651_4	0.115

	334927		CH22_FGENES.460_1	0.115
	330535	U11872	Human interleukin-8 receptor type B (IL8RB) mRNA, splice variant IL8RB1	0.856
5	328591		CH.07_hs gij5868227	0.115
	334902		CH22_FGENES.452_16	0.115
	328525		CH.07_hs gij5868482	0.115
	325870		CH.16_hs gij6682492	0.116
	337522		CH22_FGENES.819-1	0.116
10	305079	AA641329	EST singleton (not in UniGene) with exon hit	0.116
	327343		CH.01_hs gij6017017	0.116
	333918		CH22_FGENES.296_7	0.116
	333600		CH22_FGENES.213_2	0.116
	335846		CH22_FGENES.623_6	0.116
	333510		CH22_FGENES.171_4	0.116
15	327629		CH.04_hs gij5867872	0.116
	333470		CH22_FGENES.161_6	0.116
	326855		CH.20_hs gij6552460	0.116
	327008		CH.21_hs gij5867664	0.117
	337480		CH22_FGENES.795-3	0.117
20	336425		CH22_FGENES.824_10	0.117
	321964	AL079687	Hs.171065 ESTs	0.117
	335651		CH22_FGENES.590_2	0.117
	308164	AI521574	Hs.181165 eukaryotic translation elongation factor 1 alpha 1	0.117
	337927		CH22_EM:AC005500.GENSCAN.80-3	0.117
25	300341	H45095	Hs.153524 ESTs	0.117
	300154	AI245127	Hs.179331 ESTs	0.117
	306295	AA937331	EST singleton (not in UniGene) with exon hit	0.117
	329670		CH.14_p2 gij6272129	0.117
	335612		CH22_FGENES.583_6	0.117
30	307845	AI363450	EST singleton (not in UniGene) with exon hit	0.117
	330401	D28383	Human mRNA for ATP synthase B chain, 5'UTR (sequence from the 5'cap to the start codon)	0.117
	327127		CH.21_hs gij6682520	0.117
35	333843		CH22_FGENES.290_1	0.117
	331083	R17762	Hs.22292 ESTs	0.117
	329140		CH.X_hs gij6017060	0.117
	339338		CH22_BA354112.GENSCAN.27-3	0.117
	331974	AA464518	Hs.99616 ESTs	0.117
40	338631		CH22_EM:AC005500.GENSCAN.454-2	0.117
	330299		CH.06_p2 gij2905881	0.117
	330351		CH.09_p2 gij3056622	0.117
	305377	AA715714	Hs.181357 laminin receptor 1 (67kD; ribosomal protein SA)	0.117
	333106		CH22_FGENES.79_12	0.117
45	338514		CH22_EM:AC005500.GENSCAN.392-4	0.117
	327335		CH.01_hs gij5902477	0.117
	301970	AB028962	Hs.120245 KIAA1039 protein	0.118
	326339		CH.17_hs gij6056311	0.118
	330612	X15673	Hs.93174 Human endogenous retrovirus pHE.1 (ERV9)	0.118
50	334178		CH22_FGENES.350_6	0.118
	328008		CH.06_hs gij5902482	0.118
	329976		CH.16_p2 gij4878063	0.118
	320952	AA897432	Hs.130411 ESTs	0.118
	305621	AA789095	EST singleton (not in UniGene) with exon hit	0.118
55	337850		CH22_EM:AC005500.GENSCAN.34-3	0.118
	333626		CH22_FGENES.224_2	0.118
	337672		CH22_EM:AC000097.GENSCAN.67-1	0.118
	328803		CH.07_hs gij6004475	0.118
	325922		CH.16_hs gij5867122	0.118
	334489		CH22_FGENES.397_1	0.118
60	320638	R54766	Hs.101120 ESTs	0.118
	321932	AA569229	EST cluster (not in UniGene)	0.118
	336958		CH22_FGENES.367-1	0.118
	332082	AA600176	Hs.112345 ESTs	0.118
65	306004	AA889992	EST singleton (not in UniGene) with exon hit	0.118
	336803		CH22_FGENES.194-1	0.118
	309107	AI925823	EST singleton (not in UniGene) with exon hit	0.118
	336859		CH22_FGENES.293-9	0.118
	337935		CH22_EM:AC005500.GENSCAN.85-6	0.118
	326492		CH.19_hs gij5867422	0.118

	327289		CH.01_hs gi 5867481	0.119
	325818		CH.14_hs gi 6682490	0.119
	310787	AW262580	Hs.159040 ESTs	0.119
5	330028		CH.16_p2 gi 6671908	0.119
	325317		CH.11_hs gi 5866878	0.119
	335279		CH22_FGENES.523_7	0.119
	331720	AA192173	Hs.221530 ESTs	0.119
	329186		CH.X_hs gi 5868711	0.119
10	316012	AA764950	Hs.119898 ESTs	0.119
	338316		CH22_EM:AC005500.GENSCAN.304-2	0.119
	326033		CH.17_hs gi 5867178	0.119
	334745		CH22_FGENES.426_3	0.119
	333051		CH22_FGENES.73_5	0.119
15	301763	R01279	EST cluster (not in UniGene) with exon hit	0.12
	304502	AA454809	Hs.172928 collagen; type I; alpha 1	0.12
	335680		CH22_FGENES.594_5	0.12
	304678	AA548556	EST singleton (not in UniGene) with exon hit	0.12
	335441		CH22_FGENES.560_4	0.12
20	336187		CH22_FGENES.717_11	0.12
	309422	AW087175	EST singleton (not in UniGene) with exon hit	0.12
	336047		CH22_FGENES.679_9	0.12
	309651	AW195850	EST singleton (not in UniGene) with exon hit	0.12
	308547	AI695385	Hs.201903 EST	0.12
25	304443	AA399444	EST singleton (not in UniGene) with exon hit	0.12
	336245		CH22_FGENES.746_3	0.12
	302703	H72333	EST cluster (not in UniGene) with exon hit	0.12
	335690		CH22_FGENES.596_5	0.12
	328941		CH.08_hs gi 6456765	0.12
30	333873		CH22_FGENES.291_9	0.12
	317246	AW105092	Hs.155690 ESTs	0.12
	339288		CH22_BA354112.GENSCAN.16-6	0.12
	337996		CH22_EM:AC005500.GENSCAN.116-3	0.12
	333304		CH22_FGENES.137_1	0.121
35	308332	AI591235	EST singleton (not in UniGene) with exon hit	0.121
	329319		CH.X_hs gi 6381976	0.121
	302086	X57138	multiple UniGene matches	0.121
	333290		CH22_FGENES.129_2	0.121
	323825	AI793080	Hs.123525 ESTs; Weakly similar to NEUTROPHIL GELATINASE-ASSOCIATED LIPOCALIN PRECURSOR [R.norvegicus]	0.121
40	330575	U64105	Hs.252280 Rho guanine nucleotide exchange factor (GEF) 1	0.121
	305274	AA679990	Hs.181165 eukaryotic translation elongation factor 1 alpha 1	0.121
	333647		CH22_FGENES.235_2	0.121
	302251	AA333340	EST cluster (not in UniGene) with exon hit	0.121
45	329777		CH.14_p2 gi 6002090	0.121
	333155		CH22_FGENES.89_5	0.121
	326122		CH.17_hs gi 5867194	0.121
	335310		CH22_FGENES.532_3	0.121
	335453		CH22_FGENES.562_13	0.122
50	305103	AA643329	Hs.111334 ferritin; light polypeptide	0.122
	337284		CH22_FGENES.667-2	0.122
	337418		CH22_FGENES.758-4	0.122
	313073	AI963740	Hs.46826 ESTs	0.122
	303759	AW504164	EST cluster (not in UniGene) with exon hit	0.122
55	300017	M33197	AFFX control: GAPDH	0.122
	316725	AW135084	Hs.127264 ESTs	0.122
	330738	AA293153	Hs.120980 nuclear receptor co-repressor 2	0.122
	336466		CH22_FGENES.829_25	0.122
	335956		CH22_FGENES.647_3	0.122
60	315308	AA780564	Hs.189053 ESTs	0.122
	338925		CH22_DJ32110.GENSCAN.14-3	0.122
	334969		CH22_FGENES.466_2	0.122
	322050	AL137589	EST cluster (not in UniGene)	0.122
65	339084		CH22_DA59H18.GENSCAN.38-2	0.122
	338323		CH22_EM:AC005500.GENSCAN.306-2	0.122
	337003		CH22_FGENES.419-7	0.122
	325470		CH.12_hs gi 6017034	0.123
	336503		CH22_FGENES.833_10	0.123
	330786	D60374	Hs.258712 EST	0.123

	329446		CH.Y_hs gij5868886	0.123
	303326	AA229433	Hs.222634 ESTs; Moderately similar to ubiquitin-like protein / ribosomal protein S30	0.123
5	309067	AI916313	Hs.212788 EST	0.123
	317464	AA968472	Hs.130463 ESTs	0.123
	328755		CH.07_hs gij5868301	0.123
	326036		CH.17_hs gij5867178	0.123
	327208		CH.01_hs gij5867447	0.123
	326124		CH.17_hs gij5916395	0.123
10	327509		CH.02_hs gij6117815	0.123
	338398		CH22_EM:AC005500.GENSCAN.336-5	0.123
	304652	AA527782	Hs.84298 CD74 antigen (invariant polypeptide of major histocompatibility complex; class II antigen-associated)	0.123
15	335797		CH22_FGENES.612_6	0.124
	336714		CH22_FGENES.76-29	0.124
	327204		CH.01_hs gij5867447	0.124
	331881	AA430672	Hs.123778 ESTs	0.124
	306971	AI126509	EST singleton (not in UniGene) with exon hit	0.124
20	336174		CH22_FGENES.710_1	0.124
	336126		CH22_FGENES.701_13	0.124
	329129		CH.X_hs gij6588026	0.124
	303049	AW407562	EST cluster (not in UniGene) with exon hit	0.124
	335778		CH22_FGENES.607_14	0.124
25	336601		CH22_FGENES.369_2	0.124
	334340		CH22_FGENES.375_17	0.124
	337436		CH22_FGENES.767-1	0.124
	306013	AA896990	EST singleton (not in UniGene) with exon hit	0.124
	339213		CH22_FF113D11.GENSCAN.6-8	0.124
30	335355		CH22_FGENES.541_2	0.124
	336552		CH22_FGENES.841_9	0.124
	336384		CH22_FGENES.822_4	0.124
	310485	AI286202	Hs.149800 ESTs	0.125
	335840		CH22_FGENES.622_3	0.125
35	336444		CH22_FGENES.827_10	0.125
	315703	N36070	EST cluster (not in UniGene)	0.125
	327763		CH.05_hs gij5867961	0.125
	336383		CH22_FGENES.822_3	0.125
	333496		CH22_FGENES.168_6	0.125
40	328662		CH.07_hs gij6004473	0.125
	338986		CH22_DA59H18.GENSCAN.5-1	0.125
	328311		CH.07_hs gij5868371	0.125
	337241		CH22_FGENES.644-2	0.125
	336933		CH22_FGENES.350-7	0.125
45	313483	AW294432	Hs.144252 ESTs	0.125
	326116		CH.17_hs gij5867193	0.125
	330450	HG363-HT363	Epidermal Growth Factor Receptor-Related Protein	0.125
	307491	AI268539	EST singleton (not in UniGene) with exon hit	0.125
	331852	AA418988	Hs.98314 Homo sapiens mRNA; cDNA DKFZp586L0120 (from clone DKFZp586L0120)	0.125
50	330462	HG944-HT944	Dopamine Receptor D4	0.125
	304410	AA284508	EST singleton (not in UniGene) with exon hit	0.125
	336385		CH22_FGENES.822_5	0.125
	336793		CH22_FGENES.176-3	0.125
55	326243		CH.17_hs gij5867261	0.125
	327266		CH.01_hs gij5867462	0.125
	320753	AF070579	Hs.181544 Homo sapiens clone 24487 mRNA sequence	0.125
	336960		CH22_FGENES.369-5	0.125
	329667		CH.14_p2 gij6272129	0.125
60	328168		CH.06_hs gij5868071	0.125
	336534		CH22_FGENES.839_16	0.125
	339289		CH22_BA354112.GENSCAN.16-9	0.126
	309230	AI970747	EST singleton (not in UniGene) with exon hit	0.126
	339190		CH22_FF113D11.GENSCAN.1-2	0.126
	337086		CH22_FGENES.458-14	0.126
65	319233	R21054	Hs.211522 ESTs	0.126
	339396		CH22_BA232E17.GENSCAN.6-8	0.126
	331930	AA449077	Hs.179765 Homo sapiens mRNA; cDNA DKFZp586H1921 (from clone DKFZp586H1921)	0.126
	308099	AI475914	EST singleton (not in UniGene) with exon hit	0.126

	338477		CH22_EM:AC005500.GENSCAN.373-5	0.126
	334286		CH22_FGENES.369_16	0.126
	317245	AI025039	Hs.131732 ESTs	0.126
5	335249		CH22_FGENES.516_10	0.126
	333327		CH22_FGENES.138_20	0.126
	304240	AA009802	EST singleton (not in UniGene) with exon hit	0.126
	335464		CH22_FGENES.562_26	0.126
	335236		CH22_FGENES.515_8	0.126
	334154		CH22_FGENES.340_4	0.126
10	309257	AI984183	EST singleton (not in UniGene) with exon hit	0.126
	310015	AI220122	Hs.201981 ESTs; Weakly similar to breast carcinoma-associated antigen [H.sapiens]	0.126
	328280		CH.07_hs gi 5868352	0.126
15	305744	AA831819	EST singleton (not in UniGene) with exon hit	0.126
	327430		CH.02_hs gi 5867754	0.126
	328323		CH.07_hs gi 5868373	0.126
	333274		CH22_FGENES.123_2	0.126
	337193		CH22_FGENES.575-3	0.127
	334820		CH22_FGENES.437_2	0.127
20	328706		CH.07_hs gi 5868270	0.127
	331228	W67267	Hs.174911 ESTs	0.127
	307205	AI192479	EST singleton (not in UniGene) with exon hit	0.127
	337123		CH22_FGENES.519-3	0.127
	326201		CH.17_hs gi 5867216	0.127
25	335276		CH22_FGENES.523_2	0.127
	331202	T81115	Hs.191136 ESTs	0.127
	330532	U03187	Hs.121544 interleukin 12 receptor; beta 1	0.127
	321235	N49521	EST cluster (not in UniGene)	0.127
	301743	F12605	Hs.204529 ESTs; Weakly similar to reverse transcriptase [H.sapiens]	0.127
30	328175		CH.06_hs gi 5868073	0.127
	306407	AA971985	EST singleton (not in UniGene) with exon hit	0.127
	327145		CH.01_hs gi 5867548	0.127
	327649		CH.04_hs gi 5867899	0.127
	335142		CH22_FGENES.498_12	0.127
35	333909		CH22_FGENES.295_2	0.127
	330608	X04325	Hs.2679 gap junction protein; beta 1; 32kD (connexin 32; Charcot-Marie-Tooth neuropathy; X-linked)	0.127
	330158		CH.21_p2 gi 6580367	0.127
40	320153	AF064594	Hs.120360 phospholipase A2; group VI	0.127
	314407	AA098835	Hs.224432 ESTs	0.127
	333383		CH22_FGENES.143_22	0.127
	320663	AI734242	Hs.244473 ESTs	0.128
	326233		CH.17_hs gi 5867232	0.128
	326598		CH.20_hs gi 5867634	0.128
45	335174		CH22_FGENES.504_4	0.128
	319843	H29920	Hs.99486 ESTs; Weakly similar to aralar1 [H.sapiens]	0.128
	335458		CH22_FGENES.562_18	0.128
	332997		CH22_FGENES.58_4	0.128
	334188		CH22_FGENES.352_3	0.128
50	329759		CH.14_p2 gi 6048280	0.128
	330348		CH.09_p2 gi 4544475	0.128
	326958		CH.21_hs gi 6469836	0.128
	305263	AA679467	EST singleton (not in UniGene) with exon hit	0.128
55	337693		CH22_EM:AC000097.GENSCAN.78-14	0.128
	326812		CH.20_hs gi 6682504	0.128
	333237		CH22_FGENES.108_7	0.128
	333699		CH22_FGENES.250_13	0.128
	311496	AI768677	Hs.209888 ESTs; Weakly similar to phosphatidylserine synthase-2 [M.musculus]	0.128
60	336499		CH22_FGENES.833_4	0.128
	320087	AF032387	Hs.113265 small nuclear RNA activating complex; polypeptide 4; 190kD	0.128
	309989	AI184186	Hs.197813 ESTs	0.128
	301490	AW298468	Hs.250461 ESTs	0.128
	337011		CH22_FGENES.427-6	0.128
65	315052	AA876910	Hs.134427 ESTs	0.128
	301611	W22172	Hs.59038 ESTs	0.128
	336497		CH22_FGENES.833_2	0.129
	302068	Y16280	Hs.132049 endothelin type b receptor-like protein 2	0.129
	334502		CH22_FGENES.397_18	0.129

	304332	AA158884		EST singleton (not in UniGene) with exon hit	0.129
	304522	AA465405		EST singleton (not in UniGene) with exon hit	0.129
	312407	R46180	Hs.153485	ESTs	0.129
	310098	AI685841	Hs.161354	ESTs	0.129
5	301119	AF142579		EST cluster (not in UniGene) with exon hit	0.129
	309268	AI985821	Hs.62954	ferritin; heavy polypeptide 1	0.129
	330989	H42142	Hs.226396	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 19 (Dbp5; yeast; homolog)	0.129
	336949			CH22_FGENES.361-4	0.129
10	330115			CH.19_p2 gi 6015202	0.129
	339212			CH22_FF113D11.GENSCAN.6-7	0.129
	326951			CH.21_hs gi 6004446	0.129
	305165	AA662939		EST singleton (not in UniGene) with exon hit	0.129
	308238	AI559492		EST singleton (not in UniGene) with exon hit	0.129
15	337140			CH22_FGENES.537-5	0.13
	321758	U29112		EST cluster (not in UniGene)	0.13
	304619	AA515554	Hs.119598	ribosomal protein L3	0.13
	312469	AA745289	Hs.173088	ESTs	0.13
	339017			CH22_DA59H18.GENSCAN.20-6	0.13
20	330116			CH.19_p2 gi 6015202	0.13
	333312			CH22_FGENES.138_4	0.13
	338004			CH22_EM:AC005500.GENSCAN.121-1	0.13
	314141	AA232134	Hs.190028	ESTs	0.13
	300509	AI239845	Hs.128494	ESTs; Weakly similar to EG:95B7.2 [D.melanogaster]	0.13
25	338530			CH22_EM:AC005500.GENSCAN.398-11	0.13
	335968			CH22_FGENES.652_1	0.13
	314121	AI732100	Hs.187619	ESTs	0.13
	337593			CH22_C20H12.GENSCAN.6-8	0.13
	332881			CH22_FGENES.33_1	0.13
30	305836	AA858043		EST singleton (not in UniGene) with exon hit	0.13
	339059			CH22_DA59H18.GENSCAN.30-5	0.13
	305610	AA782319		EST singleton (not in UniGene) with exon hit	0.13
	305852	AA862455		EST singleton (not in UniGene) with exon hit	0.13
	327409			CH.02_hs gi 5867750	0.13
35	312751	AI613089	Hs.164178	ESTs	0.13
	308726	AI799268	Hs.209929	EST	0.13
	325961			CH.16_hs gi 5867147	0.13
	311159	AW025919	Hs.197636	ESTs	0.13
	322715	AA057230	Hs.182135	ESTs	0.13
40	336441			CH22_FGENES.827_7	0.13
	336339			CH22_FGENES.814_12	0.13
	306911	AI095365		EST singleton (not in UniGene) with exon hit	0.13
	333613			CH22_FGENES.217_8	0.13
45	338489			CH22_EM:AC005500.GENSCAN.384-17	0.131
	326904			CH.21_hs gi 5867684	0.131
	337337			CH22_FGENES.717-1	0.131
	326752			CH.20_hs gi 5867615	0.131
	303977	AW512978		EST singleton (not in UniGene) with exon hit	0.131
	301373	AA595235		EST cluster (not in UniGene) with exon hit	0.131
50	338448			CH22_EM:AC005500.GENSCAN.359-22	0.131
	333774			CH22_FGENES.272_5	0.131
	332986			CH22_FGENES.54_8	0.131
	335362			CH22_FGENES.541_12	0.131
	335896			CH22_FGENES.635_4	0.131
55	337825			CH22_EM:AC005500.GENSCAN.13-19	0.131
	325257			CH.11_hs gi 5866895	0.131
	331188	T50240	Hs.167837	ESTs	0.131
	330645	Y08302	Hs.144879	dual specificity phosphatase 9	0.131
	331760	AA292721	Hs.154434	ESTs; Weakly similar to unknown [H.sapiens]	0.131
60	322995	AA513829	Hs.29797	ribosomal protein L10	0.131
	335497			CH22_FGENES.571_5	0.131
	334824			CH22_FGENES.437_6	0.131
	319480	R06933	Hs.184221	ESTs	0.131
	334842			CH22_FGENES.439_21	0.131
65	333335			CH22_FGENES.139_4	0.131
	317252	AA905178	Hs.130124	ESTs	0.131
	329034			CH.X_hs gi 5868561	0.131
	305186	AA664230		EST singleton (not in UniGene) with exon hit	0.131
	335755			CH22_FGENES.604_4	0.131

	302143	H15270	Hs.189847	putative neuronal cell adhesion molecule	0.131
	334939			CH22_FGENES.465_3	0.131
	318994	C15110	Hs.17802	ESTs	0.131
5	334498			CH22_FGENES.397_14	0.131
	333413			CH22_FGENES.146_2	0.132
	329676			CH.14_p2 gi 6272128	0.132
	327277			CH.01_hs gi 5867473	0.132
	305022	AA627416		EST singleton (not in UniGene) with exon hit	0.132
10	336805			CH22_FGENES.196-3	0.132
	320121	T93657		EST cluster (not in UniGene)	0.132
	334761			CH22_FGENES.428_10	0.132
	339400			CH22_BA232E17.GENSCAN.7-6	0.132
	330301			CH.06_p2 gi 2905862	0.132
15	316822	AA827691	Hs.129967	ESTs; Weakly similar to neuronal thread protein	
				AD7c-NTP [H.sapiens]	0.132
	328020			CH.06_hs gi 5902482	0.132
	325327			CH.11_hs gi 5866875	0.132
	321163	AA209530		EST cluster (not in UniGene)	0.132
20	336393			CH22_FGENES.823_5	0.132
	325905			CH.16_hs gi 5867104	0.132
	305237	AA676286	Hs.2186	eukaryotic translation elongation factor 1 gamma	0.132
	339046			CH22_DA59H18.GENSCAN.28-6	0.132
	325375			CH.12_hs gi 5866920	0.132
25	333961			CH22_FGENES.304_7	0.132
	335450			CH22_FGENES.562_8	0.133
	302286	R58438		EST cluster (not in UniGene) with exon hit	0.133
	335116			CH22_FGENES.496_3	0.133
	327333			CH.01_hs gi 5902477	0.133
30	308070	AI470948		EST singleton (not in UniGene) with exon hit	0.133
	308311	AI581855		EST singleton (not in UniGene) with exon hit	0.133
	320813	AW360847	Hs.208839	ESTs	0.133
	323665	AW248307		EST cluster (not in UniGene)	0.133
	328318			CH.07_hs gi 5868373	0.133
35	320603	R51419		EST cluster (not in UniGene)	0.133
	332791			CH22_FGENES.3_1	0.133
	314976	AA524725	Hs.162108	ESTs	0.133
	303309	AL134164	Hs.224868	ESTs	0.133
	320581	R39753	Hs.170187	ESTs	0.133
40	333944			CH22_FGENES.302_2	0.133
	317992	AI733512	Hs.130901	ESTs	0.133
	330935	F02383	Hs.26492	beta-1,3-glucuronyltransferase 3 (glucuronosyltransferase I)	0.133
	336659			CH22_FGENES.36-5	0.133
	338887			CH22_DJ32110.GENSCAN.6-10	0.133
45	305273	AA679979	Hs.181165	eukaryotic translation elongation factor 1 alpha 1	0.133
	333566			CH22_FGENES.183_2	0.134
	316952	AW450033	Hs.163312	ESTs	0.134
	333818			CH22_FGENES.283_1	0.134
	328687			CH.07_hs gi 5868262	0.134
50	302879	H11802		EST cluster (not in UniGene) with exon hit	0.134
	336557			CH22_FGENES.842_2	0.134
	335222			CH22_FGENES.513_5	0.134
	338094			CH22_EM:AC005500.GENSCAN.179-3	0.134
	337384			CH22_FGENES.745-1	0.134
55	327360			CH.01_hs gi 6552411	0.134
	328132			CH.06_hs gi 5868038	0.134
	323604	AI751438	Hs.182827	ESTs; Weakly similar to !!!! ALU SUBFAMILY SQ	
				WARNING ENTRY !!!!	0.134
	337591			CH22_C20H12.GENSCAN.6-6	0.134
60	307018	AI140639		EST singleton (not in UniGene) with exon hit	0.134
	326896			CH.21_hs gi 5867680	0.134
	333479			CH22_FGENES.163_5	0.134
	337915			CH22_EM:AC005500.GENSCAN.61-3	0.134
	335110			CH22_FGENES.494_18	0.134
	333481			CH22_FGENES.163_9	0.134
65	327512			CH.02_hs gi 6117815	0.134
	300096	AW328639	Hs.83575	ESTs; Weakly similar to ZC328.3 [C.elegans]	0.134
	330163			CH.02_p2 gi 6042042	0.135
	335752			CH22_FGENES.604_1	0.135
	334857			CH22_FGENES.443_1	0.135

	301872	H84730		EST cluster (not in UniGene) with exon hit	0.135
	337529			CH22_FGENES.823-29	0.135
	335734			CH22_FGENES.601_4	0.135
	337551			CH22_FGENES.847-8	0.135
5	309078	AI920965	Hs.77961	major histocompatibility complex; class I; B	0.135
	335513			CH22_FGENES.571_28	0.135
	339078			CH22_DA59H18.GENSCAN.37-6	0.135
	321907	N56660	Hs.148722	ESTs; Weakly similar to large tumor suppressor 1 [H.sapiens]	0.135
10	337189			CH22_FGENES.571-32	0.135
	329635			CH.12_p2 gij5302817	0.135
	308601	AI719930		EST singleton (not in UniGene) with exon hit	0.135
	305020	AA627248	Hs.2064	vimentin	0.135
	333894			CH22_FGENES.293_1	0.135
15	322465	AA137152	Hs.3784	ESTs; Highly similar to phosphoserine aminotransferase [H.sapiens]	0.135
	305601	AA780975		EST singleton (not in UniGene) with exon hit	0.135
	332186	H10781	Hs.141051	ESTs; Moderately similar to !!!! ALU SUBFAMILY SB WARNING ENTRY	0.135
20	327822			CH.05_hs gij5667968	0.135
	310087	AI393914	Hs.160624	ESTs; Weakly similar to similar to CR16; SH3 domain binding protein	0.135
	328752			CH.07_hs gij5868298	0.135
	337611			CH22_C20H12.GENSCAN.19-4	0.135
25	334470			CH22_FGENES.394_1	0.136
	335115			CH22_FGENES.496_2	0.136
	328730			CH.07_hs gij5868289	0.136
	330350			CH.09_p2 gij3056622	0.136
	336971			CH22_FGENES.378-6	0.136
30	308258	AI565612		EST singleton (not in UniGene) with exon hit	0.136
	326745			CH.20_hs gij5867611	0.136
	335440			CH22_FGENES.560_3	0.136
	320257	AA330746		EST cluster (not in UniGene)	0.136
	328677			CH.07_hs gij5868256	0.136
	329731			CH.14_p2 gij6065783	0.136
35	315950	AA700553	Hs.206974	ESTs	0.136
	330049			CH.17_p2 gij4567182	0.136
	337070			CH22_FGENES.448-3	0.136
	304095	H11324	Hs.31059	EST	0.136
40	309304	AW005527	Hs.232820	EST	0.136
	333458			CH22_FGENES.157_7	0.136
	329899			CH.15_p2 gij6563505	0.136
	322202	AI275056	Hs.200133	ESTs	0.136
	333991			CH22_FGENES.310_15	0.136
45	318617	AW247252	Hs.75514	nucleoside phosphorylase	0.136
	310623	AI341586	Hs.195588	ESTs	0.136
	330489	M23323	Hs.3003	CD3E antigen; epsilon polypeptide (TIT3 complex)	0.136
	309646	AW194694		EST singleton (not in UniGene) with exon hit	0.136
	331068	R00071	Hs.191199	ESTs	0.136
50	334285			CH22_FGENES.369_15	0.136
	332178	F13689	Hs.100725	EST	0.136
	305724	AA827608		EST singleton (not in UniGene) with exon hit	0.136
	303158	AL138110	Hs.8594	Homo sapiens mRNA containing (CAG)4 repeat; clone CZ-CAG-7	0.136
	334543			CH22_FGENES.403_8	0.136
55	335384			CH22_FGENES.543_26	0.136
	336527			CH22_FGENES.839_8	0.136
	334951			CH22_FGENES.465_20	0.136
	325882			CH.16_hs gij5867087	0.137
	305134	AA653159		EST singleton (not in UniGene) with exon hit	0.137
60	307058	AI148709		EST singleton (not in UniGene) with exon hit	0.137
	331943	AA453418	Hs.178272	ESTs	0.137
	331116	R44780	Hs.22634	ESTs	0.137
	306094	AA908877		EST singleton (not in UniGene) with exon hit	0.137
	333561			CH22_FGENES.180_18	0.137
65	321439	H61962		EST cluster (not in UniGene)	0.137
	324594	AA497090		EST cluster (not in UniGene)	0.137
	337926			CH22_EM:AC005500.GENSCAN.77-4	0.137
	337353			CH22_FGENES.726-1	0.137
	331836	AA412295	Hs.104774	EST	0.137
	308981	AI873242		EST singleton (not in UniGene) with exon hit	0.137

	329424		CH.Y_hs gij5868879	0.137
	325829		CH.15_hs gij5867052	0.137
	331845	AA416863	Hs.98183 ESTs	0.137
	333854		CH22_FGENES.290_13	0.137
5	306591	AI000248	EST singleton (not in UniGene) with exon hit	0.137
	328948		CH.08_hs gij6456765	0.137
	338935		CH22_DJ32i10.GENSCAN.18-12	0.137
	325960		CH.16_hs gij5867147	0.137
	328377		CH.07_hs gij5868390	0.138
10	308851	AI829820	EST singleton (not in UniGene) with exon hit	0.138
	314620	AA424352	Hs.210586 ESTs	0.138
	337592		CH22_C20H12.GENSCAN.6-7	0.138
	338684		CH22_EM:AC005500.GENSCAN.472-3	0.138
	331800	AA400498	Hs.97543 ESTs	0.138
15	304587	AA505535	EST singleton (not in UniGene) with exon hit	0.138
	333981		CH22_FGENES.310_4	0.138
	332452	AA040369	Hs.11170 SYT interacting protein	0.138
	305752	AA835278	EST singleton (not in UniGene) with exon hit	0.138
	311947	T65554	Hs.251591 EST	0.138
20	333783		CH22_FGENES.273_5	0.138
	337406		CH22_FGENES.754-14	0.138
	327976		CH.06_hs gij5868212	0.138
	325593		CH.13_hs gij5866992	0.138
	339425		CH22_DJ579N16.GENSCAN.14-4	0.138
25	304475	AA428879	EST singleton (not in UniGene) with exon hit	0.138
	309488	AW131104	EST singleton (not in UniGene) with exon hit	0.138
	337532		CH22_FGENES.827-6	0.138
	317234	AA904448	Hs.126368 ESTs	0.138
	312261	AA854425	Hs.144455 ESTs	0.138
30	328927		CH.08_hs gij5868500	0.138
	336424		CH22_FGENES.824_9	0.138
	326667		CH.20_hs gij6552455	0.138
	325988		CH.16_hs gij5867064	0.138
	318446	AW300287	EST cluster (not in UniGene)	0.139
35	336511		CH22_FGENES.834_6	0.139
	335204		CH22_FGENES.508_13	0.139
	303244	AA147472	EST cluster (not in UniGene) with exon hit	0.139
	330870	AA115804	Hs.187593 ESTs	0.139
	329376		CH.X_hs gij5868859	0.139
40	304703	AA563898	EST singleton (not in UniGene) with exon hit	0.139
	333653		CH22_FGENES.239_2	0.139
	306799	AI051696	EST singleton (not in UniGene) with exon hit	0.139
	304872	AA595289	EST singleton (not in UniGene) with exon hit	0.139
	330812	AA013001	Hs.60563 ESTs	0.139
45	329568		CH.10_p2 gij3962490	0.139
	319210	AA253074	Hs.146261 ESTs	0.139
	334320		CH22_FGENES.374_5	0.139
	300860	AI916949	Hs.149748 ESTs; Weakly similar to weak similarity to collagens [C.elegans]	0.139
	305866	AA864533	EST singleton (not in UniGene) with exon hit	0.139
50	312943	AA984364	Hs.119064 ESTs	0.139
	330523	M99439	Hs.83958 transducin-like enhancer of split 4; homolog of Drosophila E(sp1)	0.139
	312708	AI076204	Hs.135440 ESTs	0.139
	309366	AW072970	EST singleton (not in UniGene) with exon hit	0.139
	303273	AA316069	EST cluster (not in UniGene) with exon hit	0.139
55	317484	AW274696	Hs.143921 ESTs	0.139
	333239		CH22_FGENES.111_1	0.139
	307126	AI184951	EST singleton (not in UniGene) with exon hit	0.139
	316813	AA826505	Hs.124517 ESTs	0.139
	331746	AA281365	Hs.121640 ESTs; Weakly similar to KIAA0386 [H.sapiens]	0.139
60	308558	AI700145	Hs.172182 poly(A)-binding protein; cytoplasmic 1	0.139
	310784	AW086142	Hs.159017 ESTs	0.139
	323831	AA335715	Hs.200299 ESTs	0.139
	307692	AI318342	EST singleton (not in UniGene) with exon hit	0.139
	310570	AI318327	EST cluster (not in UniGene)	0.139
65	327934		CH.06_hs gij5868184	0.139
	305232	AA670052	Hs.195188 glyceraldehyde-3-phosphate dehydrogenase	0.139
	334756		CH22_FGENES.428_5	0.139
	331938	AA451867	Hs.99255 ESTs	0.139
	301393	AI474722	Hs.150898 ESTs; Weakly similar to KIAA0644 protein [H.sapiens]	0.139

	312005	T78450	Hs.13941	ESTs	0.139
	338431			CH22_EM:AC005500.GENSCAN.351-4	0.14
	331214	T90496	Hs.16757	ESTs	0.14
5	333601			CH22_FGENES.213_4	0.14
	323481	AA278449	Hs.137429	ESTs	0.14
	336911			CH22_FGENES.344-4	0.14
	338157			CH22_EM:AC005500.GENSCAN.209-5	0.14
	327845			CH.05_hs gi 6531962	0.14
10	319109	Z45662	Hs.90797	Homo sapiens clone 23620 mRNA sequence	0.14
	334763			CH22_FGENES.428_12	0.14
	329384			CH.X_hs gi 5868869	0.14
	302996	AF054663		EST cluster (not in UniGene) with exon hit	0.14
	323751	AW452656	Hs.209824	ESTs	0.14
15	329916			CH.16_p2 gi 6223624	0.14
	301993	N49826	Hs.18602	ESTs	0.14
	338129			CH22_EM:AC005500.GENSCAN.197-2	0.14
	325704			CH.14_hs gi 5867028	0.14
	335656			CH22_FGENES.590_7	0.14
20	331673	W72366	Hs.40033	ESTs	0.14
	316807	AI018331	Hs.172444	ESTs; Highly similar to transcription regulator [M.musculus]	0.14
	310743	AW449754	Hs.158665	ESTs	0.14
	326941			CH.21_hs gi 6004446	0.14
	328809			CH.07_hs gi 5868327	0.14
25	323855	AI653164	Hs.128665	ESTs	0.14
	304705	AA564064		EST singleton (not in UniGene) with exon hit	0.14
	325666			CH.14_hs gi 6469822	0.14
	333747			CH22_FGENES.265_6	0.14
	318287	AW015616	Hs.143321	ESTs	0.141
30	332972			CH22_FGENES.51_5	0.141
	305704	AA825266		EST singleton (not in UniGene) with exon hit	0.141
	315699	AW182805	Hs.189183	ESTs; Weakly similar to Nod1 [H.sapiens]	0.141
	327296			CH.01_hs gi 5867492	0.141
	336400			CH22_FGENES.823_15	0.141
35	321033	H26214	Hs.20733	ESTs; Weakly similar to !!!! ALU SUBFAMILY SX WARNING ENTRY	0.141
	316522	AI475995	Hs.122910	ESTs	0.141
	335715			CH22_FGENES.599_15	0.141
	335959			CH22_FGENES.650_2	0.141
40	333259			CH22_FGENES.118_7	0.141
	337382			CH22_FGENES.744-8	0.141
	322346	AA227618	Hs.10882	HMG-box containing protein 1	0.141
	325378			CH.12_hs gi 5866920	0.141
	338500			CH22_EM:AC005500.GENSCAN.390-1	0.141
	338460			CH22_EM:AC005500.GENSCAN.362-5	0.141
45	315279	AW511138	Hs.256581	ESTs	0.141
	314439	AI539443	Hs.137447	ESTs	0.141
	333624			CH22_FGENES.222_3	0.141
	329237			CH.X_hs gi 5868729	0.141
50	330117			CH.19_p2 gi 6015201	0.141
	338017			CH22_EM:AC005500.GENSCAN.134-1	0.141
	337854			CH22_EM:AC005500.GENSCAN.38-12	0.142
	329984			CH.16_p2 gi 4646193	0.142
	305004	AA622328	Hs.162762	EST	0.142
55	302815	N40373		EST cluster (not in UniGene) with exon hit	0.142
	327823			CH.05_hs gi 5867968	0.142
	326753			CH.20_hs gi 5867616	0.142
	301201	AA904482	Hs.197775	ESTs	0.142
	334303			CH22_FGENES.373_6	0.142
60	326453			CH.19_hs gi 5867399	0.142
	311050	AI864581	Hs.215477	ESTs	0.142
	308740	AI802711	Hs.210337	EST; Weakly similar to aldolase A [H.sapiens]	0.142
	331003	H63959	Hs.142722	ESTs	0.142
	338010			CH22_EM:AC005500.GENSCAN.128-8	0.142
65	336326			CH22_FGENES.812_4	0.142
	318100	R44308	Hs.242302	ESTs	0.142
	320641	R55421		EST cluster (not in UniGene)	0.142
	325855			CH.16_hs gi 5867067	0.142
	330425	HG1728-HT1734		Non-Specific Cross Reacting Antigen (Gb:D90277), Alt. Splice Form 2	0.142

5	324583	AA425411	Hs.22581	ESTs	0.142
	326268			CH.17_hs gj 5867267	0.142
	331390	AA460341	Hs.45008	ESTs	0.142
	338904			CH22_DJ32110.GENSCAN.10-16	0.143
	333096			CH22_FGENES.79_1	0.143
10	331919	AA446869	Hs.119316	ESTs	0.143
	312214	AI248004	Hs.125187	ESTs	0.143
	323198	AW179174	Hs.7984	ESTs	0.143
	316107	AI204001	Hs.184014	ribosomal protein L31	0.143
	301335	AA885317	Hs.190511	ESTs	0.143
15	337392			CH22_FGENES.747-3	0.143
	325543			CH.12_hs gj 6682452	0.143
	305903	AA873085		EST singleton (not in UniGene) with exon hit	0.143
	332707	L35594	Hs.174185	phosphodiesterase I/nucleotide pyrophosphatase 2 (autotaxin)	0.143
	337913			CH22_EM:AC005500.GENSCAN.59-10	0.143
20	301436	AA961061	Hs.131696	ESTs	0.143
	335078			CH22_FGENES.486_5	0.143
	338451			CH22_EM:AC005500.GENSCAN.359-39	0.143
	302777	AJ230640		EST cluster (not in UniGene) with exon hit	0.143
	330464	J03068	Hs.78223	N-acylaminoacyl-peptide hydrolase	0.143
25	330988	H41411	Hs.33855	ESTs	0.143
	328939			CH.08_hs gj 6004481	0.143
	308015	AI440174	Hs.228907	EST; Weakly similar to GUANINE NUCLEOTIDE-BINDING PROTEIN BETA SUBUNIT-LIKE PROTEIN	0.143
				12.3 [H.sapiens]	0.143
	328504			CH.07_hs gj 5868471	0.143
30	332599	AA402891	Hs.32951	solute carrier family 29 (nucleoside transporters); member 2	0.143
	335744			CH22_FGENES.601_15	0.143
	322394	AF077208		EST cluster (not in UniGene)	0.143
	323892	AL042661		EST cluster (not in UniGene)	0.143
	318443	AI939323	Hs.157714	ESTs; Weakly similar to NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN; ALPHA-5 CHAIN PRECURSOR	0.143
35				[H.sapiens]	0.143
	336568			CH22_FGENES.843_7	0.143
	330958	H08815	Hs.159824	EST	0.143
	327672			CH.04_hs gj 5867843	0.143
	335900			CH22_FGENES.635_8	0.144
40	336044			CH22_FGENES.679_6	0.144
	318845	AI815951	Hs.33183	ESTs; Weakly similar to estrogen-responsive finger protein; efp [H.sapiens]	0.144
	333483			CH22_FGENES.165_2	0.144
	333337			CH22_FGENES.139_6	0.144
	305993	AA889197		EST singleton (not in UniGene) with exon hit	0.144
45	335719			CH22_FGENES.599_22	0.144
	325682			CH.14_hs gj 6138923	0.144
	327350			CH.01_hs gj 6249563	0.144
	339291			CH22_BA354112.GENSCAN.18-1	0.144
	326358			CH.18_hs gj 5867293	0.144
50	330316			CH.08_p2 gj 6007576	0.144
	308150	AI499346	Hs.174131	ribosomal protein L6	0.144
	338065			CH22_EM:AC005500.GENSCAN.164-1	0.144
	339009			CH22_DA59H18.GENSCAN.18-7	0.144
	327776			CH.05_hs gj 5867964	0.145
55	336664			CH22_FGENES.41-8	0.145
	321921	AF070619		EST cluster (not in UniGene)	0.145
	319346	T70147	Hs.12024	ESTs	0.145
	304265	AA062892		EST singleton (not in UniGene) with exon hit	0.145
	303618	Z45986	Hs.250178	copine II	0.145
60	327498			CH.02_hs gj 6017023	0.145
	335227			CH22_FGENES.513_13	0.145
	339022			CH22_DA59H18.GENSCAN.22-1	0.145
	302597	H55661	Hs.33026	ESTs; Weakly similar to similar to Enterococcus faecalis TRAB [C.elegans]	0.145
			Hs.201811	EST	0.145
65	308550	AI697008	Hs.156015	Homo sapiens chromosome 19; cosmid R29381	0.145
	302175	AA262760		EST cluster (not in UniGene) with exon hit	0.145
	303252	AA156760		CH22_FGENES.757-2	0.145
	337414			EST cluster (not in UniGene)	0.145
	310382	AI734009		CH.X_hs gj 5868806	0.145
	329333				0.145

	336857		CH22_FGENES.291-7	0.145
	332565	AA234896	Hs.25272 E1A binding protein p300	0.145
	318634	AI928098	Hs.156832 ESTs	0.145
5	336318		CH22_FGENES.801_1	0.145
	310960	AI923551	Hs.170843 ESTs	0.145
	335346		CH22_FGENES.537_2	0.145
	331196	T65416	Hs.12826 ESTs	0.145
	337607		CH22_C20H12.GENSCAN.17-3	0.146
10	331206	T84096	Hs.15284 ESTs	0.146
	301793	T80698	EST cluster (not in UniGene) with exon hit	0.146
	319590	AA210878	EST cluster (not in UniGene)	0.146
	311394	AI695374	Hs.256231 ESTs	0.146
	324773	AA632554	Hs.163401 ESTs	0.146
	324841	AI142359	Hs.155316 ESTs	0.146
15	332260	N70088	Hs.138467 ESTs	0.146
	329276		CH.X_hs gil5868762	0.146
	335887		CH22_FGENES.633_1	0.146
	338294		CH22_EM:AC005500.GENSCAN.297-1	0.146
20	336993		CH22_FGENES.409-4	0.146
	334135		CH22_FGENES.336_2	0.146
	326251		CH.17_hs gil5867263	0.146
	337396		CH22_FGENES.749-1	0.146
	339167		CH22_DA59H18.GENSCAN.69-8	0.146
25	316838	AW135418	Hs.161210 ESTs	0.146
	325313		CH.11_hs gil5866865	0.146
	331047	N66918	Hs.32205 ESTs	0.146
	323915	AL043362	EST cluster (not in UniGene)	0.146
	302747	AF062275	EST cluster (not in UniGene) with exon hit	0.146
30	306317	AA947909	EST singleton (not in UniGene) with exon hit	0.146
	334399		CH22_FGENES.382_5	0.146
	326472		CH.19_hs gil5867404	0.146
	333061		CH22_FGENES.75_4	0.146
	337072		CH22_FGENES.448-5	0.146
35	334328		CH22_FGENES.375_5	0.146
	327039		CH.21_hs gil6531965	0.146
	325576		CH.12_hs gil6552443	0.147
	315935	AI075804	Hs.132660 ESTs	0.147
	319638	AA323758	EST cluster (not in UniGene)	0.147
40	334501		CH22_FGENES.397_17	0.147
	338238		CH22_EM:AC005500.GENSCAN.264-4	0.147
	308636	AI744063	EST singleton (not in UniGene) with exon hit	0.147
	336567		CH22_FGENES.843_6	0.147
	335819		CH22_FGENES.619_2	0.147
45	336950		CH22_FGENES.361-8	0.147
	307055	AI148477	EST singleton (not in UniGene) with exon hit	0.147
	315134	AW504854	Hs.126714 ESTs	0.147
	335834		CH22_FGENES.621_1	0.147
	327870		CH.06_hs gil5868131	0.147
50	323802	AA332011	Hs.250138 protein phosphatase 2C; magnesium-dependent; catalytic subunit	0.147
	329412		CH.X_hs gil6682553	0.147
	323791	AA333068	EST cluster (not in UniGene)	0.147
	324126	AA385315	EST cluster (not in UniGene)	0.147
	327865		CH.06_hs gil5868130	0.147
55	333445		CH22_FGENES.154_2	0.147
	321302	AA021351	Hs.158497 KIAA0724 gene product	0.147
	336744		CH22_FGENES.118-9	0.147
	323731	AA323414	EST cluster (not in UniGene)	0.148
	320289	H07989	EST cluster (not in UniGene)	0.148
60	305488	AA749000	EST singleton (not in UniGene) with exon hit	0.148
	305592	AA780594	Hs.62954 ferritin; heavy polypeptide 1	0.148
	304094	H11295	EST singleton (not in UniGene) with exon hit	0.148
	325040	AW296368	EST cluster (not in UniGene)	0.148
	339034		CH22_DA59H18.GENSCAN.26-2	0.148
65	334504		CH22_FGENES.398_2	0.148
	334778		CH22_FGENES.431_2	0.148
	320148	U77494	Hs.119687 RAN binding protein 8	0.148
	303584	AW173759	Hs.203401 ESTs	0.148
	325826		CH.15_hs gil5867048	0.148
	331192	T55182	Hs.152571 ESTs; Highly similar to IGF-II mRNA-binding protein 2 [H.sapiens]	0.148

	325785		CH.14_hs gij6381957	0.148
	333166		CH22_FGENES.91_8	0.148
	336548		CH22_FGENES.841_5	0.148
	337552		CH22_C4G1.GENSCAN.1-4	0.148
5	331775	AA382742	Hs.97151 EST	0.148
	338936		CH22_DJ32110.GENSCAN.19-6	0.148
	331869	AA428554	Hs.104894 ESTs; Weakly similar to fibronectin precursor [H.sapiens]	0.148
	332865		CH22_FGENES.28_5	0.148
10	328663		CH.07_hs gij6004473	0.148
	328436		CH.07_hs gij5868417	0.148
	311158	AI634864	Hs.250789 ESTs; Highly similar to similar to NEDD-4 [H.sapiens]	0.148
	336942		CH22_FGENES.354-2	0.148
	302262	R53169	Hs.246091 ESTs	0.149
15	333296		CH22_FGENES.132_3	0.149
	333365		CH22_FGENES.142_2	0.149
	311706	AW452392	Hs.252854 ESTs	0.149
	337109		CH22_FGENES.489-2	0.149
	315062	AW173300	Hs.190201 ESTs	0.149
20	333454		CH22_FGENES.157_3	0.149
	334784		CH22_FGENES.432_9	0.149
	333255		CH22_FGENES.118_3	0.149
	337518		CH22_FGENES.814-7	0.149
	320651	AA489268	EST cluster (not in UniGene)	0.149
25	323437	AA287567	EST cluster (not in UniGene)	0.149
	328761		CH.07_hs gij5868302	0.149
	328787		CH.07_hs gij5868309	0.149
	335261		CH22_FGENES.520_2	0.149
	300827	R16689	Hs.106004 ESTs	0.149
30	339263		CH22_BA354112.GENSCAN.10-1	0.149
	337412		CH22_FGENES.756-6	0.149
	334414		CH22_FGENES.384_1	0.149
	332931		CH22_FGENES.38_5	0.149
	310801	AW270980	Hs.106346 novel centrosomal protein RanBPM	0.149
35	305216	AA669056	EST singleton (not in UniGene) with exon hit	0.149
	314779	AA470122	Hs.190261 ESTs	0.149
	338414		CH22_EM:AC005500.GENSCAN.341-27	0.149
	303342	AW247361	EST cluster (not in UniGene) with exon hit	0.149
	337509		CH22_FGENES.806-4	0.149
40	306631	AI001149	EST singleton (not in UniGene) with exon hit	0.149
	302533	L36149	Hs.248116 chemokine (C motif) XC receptor 1	0.149
	336536		CH22_FGENES.839_18	0.149
	324666	T32458	Hs.14285 ESTs	0.149
	310173	AI767433	Hs.170013 ESTs	0.149
45	333595		CH22_FGENES.211_2	0.149
	335975		CH22_FGENES.652_9	0.15
	306654	AI003654	EST singleton (not in UniGene) with exon hit	0.15
	335025		CH22_FGENES.475_3	0.15
	328711		CH.07_hs gij5868271	0.15
50	328274		CH.07_hs gij5868219	0.15
	325505		CH.12_hs gij6682451	0.15
	329641		CH.14_p2 gij6468233	0.15
	304955	AA613504	EST singleton (not in UniGene) with exon hit	0.15
	339103		CH22_DA59H18.GENSCAN.44-10	0.15
	329636		CH.12_p2 gij5302817	0.15
55	310118	AI203293	Hs.157489 ESTs	0.15
	326056		CH.17_hs gij5867184	0.15
	303773	AA769074	EST cluster (not in UniGene) with exon hit	0.15
	303153	U09759	Hs.8325 mitogen-activated protein kinase 9	0.15

TABLE 13A shows the accession numbers for those primekeys lacking unigeneID's for Table 13. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey:	Unique Eos probeset identifier number	
CAT number:	Gene cluster number	
Accession:	Genbank accession numbers	
<hr/>		
Pkey	CAT number	Accession
322050	24275_1	AL137589 AA423949 BE222949 BE222694 AI199615 AW873116 AI277950 AW044290 AW630096
321439	1599424_1	H61962 W01567 N75711
321666	13653_22	BE259906 AA232518 AA013359 AL035788 AW160822 BE387134 BE002954 BE391839 AW161565 AI878841 BE616458 BE409981 BE387308 BE297436 BE315536 AA206924 R12012 AA214169 BE312812 BE387093 H11710 BE312009 BE260569 AA343566 AA219526 R34757 AA219749 BE336733 AA219751 AW411099 AA232408 BE018716 BE398089 AA206253 AA053487 AA114224 AV655868 AW732566 BE394087 AW732574 AA313442 BE336875 AA070548 BE259840 BE019828 AW732341 AA299916 BE019253 BE018238 BE387109 AA232304 BE255589 AW732585 AA181436 AA308777 AA075802 AW732521 AA314526 AA226747 BE409513 AA206168 BE388292 BE298782 BE387086 AA305310 AV652723 AA314918 BE615510 AW951763 BE398104 BE385195 BE407165 BE391336 BE390187 BE389189 BE540650 BE249884 BE385985 BE274245 BE391124 BE260080 AA182600 BE512821 BE390090 BE279398 BE279589 BE263454 BE515194 BE293569 BE272531 BE388814 BE384659 BE271685 BE561043 BE278449 BE302572 AW239076 AI750583 AA376179 AA112632 BE266324 BE266614 R13105 AA132286 BE296305 AI220355 AA205606 AA219527 AA219519 AW804310 AA083286 BE171208 T19693 AA338328 BE185868 AA903024 T92162 AA330119 BE410404 BE314668
300088	622937_1	AW576245 BE207878 AW299993 AI199558 AI285442 AW299994 AW394242 AW394184
322303	704603_1	AI357412 AI870708 AI590539 W07459
322394	27492_1	AW068287 AA310079 BE336702 AA356318 AA306059 AA346785 AW402633 AA311210 AW402909 N76879 AW402913 AW401920 AA321636 AA354474 C17297 C16938 AA311774 M29871 NM_002872 Z82188 AW405674 H94176 R89281 AA214723 AI014482 AW949347 T27749 AW804226 AW796964 AW404581 AF077208 NM_014029 W68830 W79652 AA353375 AW575218 AA552192 AA521232 AA702695 AA033975 AW407827 AA829948 N94402 AW628604 AI523308 N57605 AA641662 H42477 N52784 AI753478 AA768493 AA845729 W47391 N55270 AI090117 R89282 BE206172 AA076650 AA595650 AI218931 BE049397 AI433110 W74114 H94277 AI358627 AI085221 AI862818 AA835967 AW103905 AI640644 AA835507 AA856887 AA694392 AW337542 AI524410 BE045500 AI440060 AI358801 AW028238 AW205248 AI718264 R48618 AA357358 AI695002 AA897549 AW081065 AI433360 AI810783 AI620963 Z82188 AA360224 U29112 AI656540 AI364875 AI656246 AI990940
321758	44275_1	AA169345 AI762857 AI949997 AI809601 AI681948 AI221079 AW167404 AI347614 AI611090 AI023472 AI347683 AI027467
323109	155498_1	AW591788 AI380665 AA835735 AA836654 AI244028 AW193159 AI500112 AI918722 AI738693 AI702308 AA805365 AI766842
322533	38937_1	T59538 T59589 T59598 T59542 AF147374
321921	34680_1	AF070619 R20302 T80358
321927	21620_1	AJ223366 BE305086 AW820106 AA621983 BE305208 AI738475 AI380189 AW590847 AI127232 AA622706 AI380858 AA621975 AI587036 AA665743 AW204003 AI692234 AI002242 AI692219 AW137282 AW268783 AW295910 AI380815 AW301462 AI318288 AI318575 AI318117 AI345591 AI249650 AI246934 AI246864 AI246971 AW268311 AI249654 BE041907 AW732776
321932	265316_1	N72324 N52825 W19526 BE143464 AA376060
306971	14694_7	M83667 NM_005195 S63168 M83667 AW068039 AW630649 AI338577 AI018125 AI269878 AW242440 AI887823 AI342581 BE222416 AI582847 AI651011 AI660815 AI699574 BE550201 AI926996 AW665855 AI827752 AI761857 BE328168 BE222451 AI762201 AW000929 AW007207 BE042962 BE551843 BE465373 AI279179 AI949945 BE551862 AW051667 BE328076 BE222296 AW007229 AW772332 AI279801 AI934526 AI631938 AI770103 BE041412 AI417900 AI692655 AI869943 AW270119 AI431739 AI703347 AW770568 AW025473 AI701497 AI128026 BE328147 AW203980 BE046793 AW087704 AI674597 AI650732 AI813691 AI472092 AI695224 AI241217 AW207746 AI206840 AI271362 AI631788 AI911883 AI914619 AI380585 AI767501 AI823759 AI564116 AI190991 AI377369 AI814122 AI221623 AI354793 AI081988 AI391740 AI337435 BE467366 AI824347 AI565325 AI280038 AI640455 AI819744 BE467803 BE327524 AI149402 AI313187 BE219684 AW611948 AW665821 AI091260 AW044492 BE220366 AW025381 AW183264 AI694865 AI498474 AI129780 AI202028 AI566792 BE220659 AI928040 AI830696 AI493021 AW612488 AI913152 BE042965 AI631837 AI693873 AI498925 AI768668 AI401544 BE327023 AI693383 AI769874 AI744003 AW082273 AI686501 AI798177 AI985196 AI090033 AI432342 AI689918 AI638308 BE468080 BE219588 AI912119 BE219787 AW005392 BE326564 AI589039 AI860187 AI758143 AI338168 AI702936 BE221985 AI498727 AI918196 AI279735 AW771497 AI860133 AW237834 AW661759 AW028111 BE503416 AI360180 AW611715 AI871777 BE045447 BE326444 AI266547 AI800237 AI823315 AI478368 AI264281 AI675841 AI690041

5 AI498018 AI554124 AI239893 AI864054 AI280099 AI192815 AI620465 AI080201 AW002057 BE500986 AI341131 AI818991
 AI566137 AI123403 BE219192 AW183844 AI499842 AW137971 AW138720 AW015526 AW138160 AW243163 AW138705
 AW139927 AW140006 AW138810 AW137450 AW206970 AW135419 AW205974 AA043494 BE465106 AW139955 AI741112
 BE326942 AA043506 AI079957 AI942432 AI392902 AI097047 AI470599 AA514553 AA984008 N47949 AI654114 AA884832
 AI796752 AI765290 AI301155 AW470358 BE222764 AI823569 AI651188 AI692695 AI476643 BE504307 AI767573 BE219719
 AI932249 AW467075 AI913633 BE221966 AI091025 AA969215 AI799810 AA931170 BE048559 AI809606 AI138614
 AI739456 AI674605 AW772068 AI089286 AI625787 AI263418 AW008638 AI928389 AW628997 AI470010 AI914168 AI760003
 AI203050 AI334069 AI694788 BE045337 AI948659 AI912982 AI867131 AI192102 AI767583 AI347518 AI566005 AI625884
 10 AI215888 AI633904 AW182265 AW614357 AI128030 AI343685 AI914283 AI985003 AI823578 AI493053 AI380285 AI633895
 AI267880 AI538162 AI991552 BE219479 BE219296 AI302178 AW779296 AI913805 AI631644 AI566772 AI985498 AI942289
 AI935659 AI339092 AI247432 AI686472 AI766886 AI017228 AI333272 AW301668 AI972218 AW082027 AI632974 AI474761
 AI766127 AW236578 AW000966 AI870734 AI222399 AI871249 AI703448 BE464210 AI768037 AI871585 AI767871 AI738757
 AI220732 AI681633 AI768783 AI684463 AI307339 AI263203 AW665264 BE463969 AI768786 AI439118 AI127913 BE218324
 AI672342 BE220052 AI796163 AI221662 AW197672 AW025300 AI769681 AW612448 BE219757 AW072420 AI669980
 15 AI830418 AW204353 AA047011 AA913868 AI739146 AI669954 AW470507 AW614835 AW302151 AW772372 AI762427
 AW339902 AW303370 BE464775 AW299818 AW236072 AW195060 AW274737 AW263062 AW183846 AI868894 AW300493
 AW172509 AW516876 AW593773 AW299474 AW303546 AI817323 AI823624 AI694005 AI934589 AI343479 AI861825
 AI962726 AI765845 AW080318 AI640227 AI763042 AI768903 AW235386 AA738489 AW341293 AA588585 BE221732
 AI914179 AW611669 AI572789 AW194735 AW236122 AW236007 AW612789 AW197501 AW195046 AI797145 AI864423
 20 AI458934 AI342848 AI693227 AI912642 AI689993 AA932572 AA740269 AW470392 AW086020 AI221701 T69326 T70461
 AI765579 AI338263 AI431721 AI394249 AI186462 AI823571 AI953665 AI497954 AI761057 AI678228 AI640302 AI948742
 AA594626 AA883155 AI972682 AI804774 AI300407 AI433524 AA897341 AI401175 AI291071 AA021213 AI126509 AI948955
 AI218835 AA903938 AA502610 AI498320 AA584267 AA935285 AI476253 AA489658 AA975053 AA715326 AA557139
 25 AA126417 AA971455 AA557319 AI499738 AA911438 AI913637 AA494506 N90793 AI990724 AA131667 AA128164
 AA046840 AI262557 AA131729 AA594926 T59467 AA436907 AA044630 AI589177 AI279237 AI880498 AI431822 AA708934
 AW612558 AI634069 W03610 AI192272 BE550862 AI400879 AA708507 AI128003 AI375308 AI271423 AI199552 AA125977
 AI366498 AA458662 AI694382 AA044627 AI636263 AI796270 T90146 AW014724 AI870812 AI948781 AA369965 AI094721
 AW271817 AI262898 AI244680 T69252 AI934148 AA046357 W19109 AA028157 AW021924 AA253491 AI189397 AI934388
 30 D58282 W21323 W24288 AI682972 AA293683 AA284566 AV659511 AA434184 H87089 AA040038 N57464 AA343709
 AW805815 R89837
 301119 33384_1 BE621320 BE266806 BE276582 AW516729 AF142579 AW451687 AK000069 AA325236 BE168997 W73105 AA715365
 BE278873 AA808894 AA386371 AW517942 AW750993 BE140314 BE392384 BE621757 AA318192 BE548173 AW152607
 AW166898 AA352215 AW841506 T59802 AF147378 AA335719 AW956069 T59668 AA826362 AI961329 AI290469
 35 AW197375 AI805651 AA160748 AA581089 AI968889 AA581100 AA501478 AI621069 AA468534 AA503715 AA658457
 AI144504 BE387827 AI159880
 AW177009 AI381610
 324019 262792_1 AW287567 AA252404 AW967735 AA287568 AA761222 AA865644 AA831245
 323437 189513_1 BE514807 R43224 AI363450 AA450226 AF030942
 307845 19804_10 AA385315 AI627453 AI050695 AI348281
 324126 272259_1 AI340462 AI583268 AA079086 AI950777 AI301866 AI925108 AW876954 AW877000 AA525418 AA888549 AI934220
 40 309101 7570_1 AW380220 AA804858 AI927576 T61151 AW384053 BE391691 AA533856 AA248400 T48202 N57156 R68346 R26020
 AL050332 W30806 H61369 AA092592 AA230324 BE271217 AW372903 T48772 AA358002 AA094302 AA559856 AW373308
 AW373315 AW373297 AW373311 AW373314 AW373309 AW877055 AW770140 AW379805 AI581609 AW364144 AA078921
 AA715432 AA654210 AI004899 AA602209 W47464 AA506588 R26822 AU076528 AI535743 AI535704 AI535681
 45 315703 119175_1 AA402307 D60405 D61237 D59891 AW964877 AA325215 AI459739 N36070 N25658 AA083684 AW293368 AI761958
 AI741205 AI693175 AW873603 AI143269 AI187124 N25199 H19323 AI650842 AW316825 AA083842 AA935650 AW298404
 AI472001 AI648568 R17676 R41625 AI123237 R17677 AI206866 F36920 AI654713 F34084 AA618029 AI915139 AW275194
 AW514577 D80420 AW149850 Z40953 AI867861 AA927547 AA974344 AI825793 AI635565 AI652157 BE504748 AW295759
 F16800 AW839796 F01781 AA909730 AA984010
 50 AA595235 AW973839 T03040
 AW248307 AA313452 AW951927 AA355961 BE566080
 AI702835 AI758919 AI685405 AI952108 AI299207 AI400767 AW105389 AI952710 AA845312 AI784118 AI537315
 X57138 NM_003514 Z98744 BE253911 BE256314 AI095013 AI138475
 55 AA323414 AW664013 AI809377 AI276041 AW296883 AI798340
 AA333068 AA331863 AA331838 AW962531 AA331442
 AW296368 AA247632 AK002030 R15304 T08775 AW975664 AI186801 AA730688 AW190918 AI141176 AW513211
 AI275071 AA988601 BE042933 BE045713 AW087176
 AA464018 AA464079 AA468142
 60 AA846318 W15478 AL042661
 AW131104 BE246610
 AA333340 AW955834 R49755 U33428
 R58438 AA358612
 AL043362 AA350031 AW751972 BE549118
 AA497090 AI351879 AI350914
 65 AI815981 AF287269 BE260960 BE263991 AA311733 F12145 F07345 Z43604 T29948 H64102 Z43611 T35364 N40667
 AI909783 AW751045 AA160594 AI816064 AI307240 AI951554 AA641031 AA293045 AI942492 AI687077 R78689 H12368
 AA894728 AI124930 AI423498 AA777559 AA614585 AW071822 T66288 AI418558 H21480 AI335011 AI051728 AA293436
 AW302233 AW188628 N26393 AI076557 AI311022 AW451505 H62593 Z39666 H12315 AI761351 AI364142 F02935
 AW571491 T35366 AI240745 H64151 AA503793 AA831948 AI627686 AI761531 F03591 F09782

5	301763	1688575_1	R01279 R05896 T86522
	301780	18597_2	R05735 BE349600 R37388 N79751 R10115 AA702039 AA836147 AA505716 AI049661 AI499239 R54072 AI023394
10			AA827710 W60285 W500038 AI884786 AA827191 AA810075 AW005088 R70248 AI858560 AW078678 AA631306 H52839
			AW085835 AI656182 AA737178 AW136923 AA281028 AA570316 AA722871 AA362737 AI217268 BE242373 R01113
15			AA628946 AI394527 AW402308 AI361110 AI917585 T99639 AA805326 N44577 AI394021 AW403385 T23949 AI497766
			T96602 AA834947 AI693908 Z33450 T92127 BE541896 AI933301 BE251540 BE252269 N50968 AI695531 AW575523
20			AW296889 N93796 N89924 AI361804 AI085251 AA810694 BE303011 AA743784 R13478 AA358771 AA325294 AW964880
			BE258953 R54116 AW881039 AW602593
25	301793	239325_1	BE265837 AA340632 T96304 T96075 T72780 H51978 R09868
	303049	102592_1	AW408042 AW407562 BE172835 BE396893 BE269184 AA045741 BE004187 AW751261 W74283
30	301863	19477_1	BE263301 AI418863 NM_005194 X52560 AW328683 BE298869 D63161
	301872	27494_4	H84730 T73262
35	301893	6561_1	T80334 BE292758 AK000854 H16996 BE253691 R88508 AA357663 AW955288 AW579550 N98864 AA595201 AI742967
			AA602658 AI091433 AA813367 AI983217 AW298007 AI628490 AI708037 AI560654 AI032983 R88509 R38972 AI687783
40			AI560153 AW874581 N69891 AA993617 H51180 AI269042 AI281358 AW591213 AI017724 AI262859 H16997 R38991
			AI804355 AI868988 AI669525 AW023081 AL047848
45	310382	653318_1	AI734009 AI263076 AW272255 AI792912
	303181	74060_2	AA452366 AA351338 BE262590 BE262591 AA074050 AA389667 BE161346
50	302569	17513_2	AC004472 BE312721 BE273942 F11928 T65358 BE612432 BE261576 BE179884
	324893	4670_1	AA324119 AW246199 BE395368 BE261676 BE382334 BE394701 BE304548 T31940 BE398128 BE398019 BE296693
55			BE379564 BE269460 BE397065 Z42029 BE305028 AK000549 BE536182 BE314372 AW393349 T50987 AA069735
			BE386997 AW381699 T51050 W95025 AA477678 AA348306 AW956831 AW062919 AL040397 BE305160 AA315419
60			AW249929 AA295944 AI635946 AI870259 AI951125 AW028250 AI885184 AW873113 AI077544 AW025091 AI817594
			AI401718 AW008245 AI499064 AA599687 AI016890 AA765638 W93340 AA588708 AW519173 R51917 AA676778 AI084871
65			AA687684 AI860840 AI811921 AW514730 AA477561 N78845 AA779894 AA778559 AI968953 T16188 T32828 AA991426
			AI474472 AI473542 AI828972 AW247906 AA977415 AW591489 AA876008 AW191893 AA074278 AW874099 Z40196
70	303244	9334_1	AW083615 F01544 T55984 AI290413 AI972167 AI365049 T36028 AI042568 BE560076 W17119 AA196376 T47999 R54309
			AK001269 AL354613 AA147472 AA490803 BE207628 AW816113 AA085574 AW503392 AA299910 AW750305 BE079539
75			BE079484 BE512838 AK001593 AW968772 AW967440 AW206280 AA251270 AI627886 AA303599 AA147473 BE206616
			AA490611 AA715039 AW590866 AW590447 AI864512 AA204731 AA894490 BE001136 AA612785 AA237035 AA149960
80			Z44257 R12986 AA448446 AI734041 AA422167 BE220551 R66041 R32927 R32942 AA258773 AW386142 R53730 N54624
			AW880296 AA253485 AW954441 H98989 AW614348 AI654838 AA779793 AW237213 N66635 AI186812 AA947479
85			BE158011 AI859480 AW805579 N52010 AA806305 AI628445 AW270990 AA778165 AA149949 AI650728 AA749108
			AA687257 AI261661 AA747442 AA481351 AA206339 AA903407 AW473306 AI688930 AA262281 AA448310 AA748820
90			AI347430 BE465692 R32839 AW510564 AA436408 AA257971 AA253362 AA938330 AA513150 AA976840 AA687117
			AI281547 AA046243 R32825 AI631554 AW139818 AI244536 R52946 AW235443 R40183 AA299909 AA811958 AI302918
95			Z40213 BE158047 BE158060 AA767245 AW748159 AW500735 AA094074
			AW393348 AW393350 AW386713 AW384705
100	303252	149690_3	AA316069 BE274224 AL120803 BE170052 BE170039 AI906340 BE091310 AA491506 AW836675 AW863111
	303273	67758_1	AW973784 AW843642 AA557573 AA578088 AI125161 AA349349 AI372794 BE312586 BE312777 T32148 AW239077
105	302640	21194_1	AI905357 Z42685 AW298772 R18578 AA780425 AA325971 AI372793 R10658 AA295021 AW885349 AW885288 BE271987
			AW366519 AA349350 AA233207 R88464 AA434299 R02058 R00019 R54563 Z44886 R20150 AW368328 AW368321
110			AW802152 W79803 H12809 AA028951 AW367382 AA295247 H46355 AA905620 R54564 H12765 AW950608 AA028952
			AA366908 AI085652 R43207 R77954 AI672848 T28547 AA427734 AA572853 AA769934 AI242108 R00020 R02059 R10659
115			AI185270 AI041890 NM_000080 X66403 F03854 AI652442 AI766431 AA976913 AI989882 AA471024 AI802727 AI824112
			F02169 AI890843 BE250876 BE252859 AL157418 R78326
120	303342	189722_1	AW068570 AW247361 AA252638 AI751982 BE260758 BE293073 AW293303
	302703	7075_1	AB040951 AK002094 AA676593 W44644 N42376 Z45942 AW841844 BE541378 AA358274 AA213391 T88771 NM_015493
125			AL117489 N88248 N31714 N36273 N31721 AW576263 AA449380 AI366135 AA551576 AW149789 Z17418 AW474331
			AA056181 AI753611 AA046428 AA488007 AA300764 N44732 AA377697 AA346752 AA485787 AA894546 AA115295
130			AA299914 R88096 AA367342 AW884666 W84522 AA426325 AI983849 AA873315 AA873307 AI355170 AA534678
			AA969227 AI127202 AW083323 AI338244 AW020877 AA780019 N33426 AW069314 N63079 AI926527 AA115270
135			AW886601 AI357402 AA599312 AI460358 AI926969 AA429402 N33197 AW886733 R88205 N52803 AW021988 AA213392
			BE139656 AI142383 AA427844 AA954743 AA233622 AW073382 AA426326 AA493560 AA425133 N24819 AI419516
140			AI571515 AI147373 AI628677 AI214877 AA992123 H71599 AA029095 AA622262 AW117398 AW275286 AI911337
			AA864950 T94173 AI475634 AI701411 AI287696 T94091 AA505746 AI184310 AI350967 AI083596 W74274 AI954381
145			AI832767 AI368443 AA195578 AW874416 AI005421 AW014339 AA908660 AI350791 AW241382 AI473104 AI275186
			AA515528 AA194897 AA782901 AW069414 F20248 AA426011 AI305169 AI832109 AI570082 AW072984 AI492474
150			AA919076 AL049024 W79889 N42400 AA625435 AW963887 AA233420
			AW779971 AW300287 AW152002 AW069505 AI866447 AI298231 AI146920 AI692267 AI872876
155	318446	604736_1	BE397032 AJ292529 N40373 N34073 AA321112 AW959902 AA258103 AW860213 BE549059 BE296027 BE296657
	302815	42200_1	AA300789 AI971491 AW513665 AA909530 AI951045 AW058103 AI971506 AI061239 AA600054 AI000807 AA969975
160			AA281492 AW593654 AA321111 AW298633 AI278754 AI863862 AI285506 AA989727 F33114 T16079 AI762625 AI492103
			AW770346 AW026768 AI468710 AI499987 AA310412 AA622784 AA642297 AI866427
165			H11802 T66097 AF042831
			R42185 AW939055 T30280 Z43366 R54166
170	318540	1018709_2	AA938905 AA574056 AA714466 AI805592 AI123431 AA229723 AA620759 AI004450 AW299820 AI949299 AW874308
	302928	22118_1	AA626037 A974112 AA931563 AF073924 AA995769 AI766441 AI367730 AI081342 AA235800 AA235801 AI138970

AA719797 AA759343 X89673 AA759344 AA312909 X87825 Y10529 AC006271 AA758739 BE501015 AA909905 AF065857
AC006271 AA970044
304332 28696_41
5
10
15
20
25
30
35
40
45
50
55
60
65
AA158883 AA171835 AA187049 AA143546 BE299538 BE614280 BE621705 BE299684 BE619550 BE613099 BE619558
BE514331 BE617716 BE612920 BE615742 BE258739 BE621539 AI434511 BE546696 BE614324 BE379359 BE250106
BE250681 BE299592 BE300272 BE616805 BE397385 BE562024 BE271246 BE250556 BE280311 BE561995 BE618755
BE276126 BE546275 BE311547 BE262155 BE281082 BE513087 BE546891 BE514289 BE397389 BE267442 BE545455
BE614483 BE293447 BE270710 BE281071 BE267458 BE542095 BE262701 BE513634 BE548116 BE299546 BE619604
BE512885 BE616638 BE266173 BE258933 BE259710 BE268569 BE563861 BE614871 BE537509 BE250108 BE515323
BE538868 BE250081 BE277706 BE410127 BE619445 BE250753 BE304969 BE616348 BE546878 BE544962 BE410346
BE267256 X17206 NM_002952 BE304541 BE619171 BE259655 BE549186 BE314944 BE613101 BE378069 BE621110
BE542752 BE257029 BE531315 BE619306 BE267328 BE259439 BE297093 BE280651 BE407684 BE250201 BE312819
BE535432 BE279917 BE312626 BE531118 BE378744 BE275370 BE250195 BE409980 BE274432 BE266637 BE279321
BE622382 BE280232 BE263816 BE378977 BE300145 BE250204 BE547609 BE264377 BE266688 BE259746 BE260829
BE619517 BE388097 BE264025 BE618945 BE614758 BE312249 BE294359 BE531121 BE622300 BE615109 BE544354
BE614998 BE393239 BE297520 BE393221 BE278818 BE279309 BE265476 BE618772 BE615185 BE265144 BE249837
BE312230 BE407843 BE253884 BE407645 BE615804 BE619058 BE559512 BE383249 BE613497 BE294351 BE295062
BE622385 BE390654 BE535438 BE563186 BE396374 BE270842 BE386110 BE260368 BE250186 BE265875 BE537229
BE253369 BE256997 BE269482 BE264959 BE279072 AA662160 BE280733 AA858428 BE561308 BE267285 BE561422
BE563181 BE304614 BE295437 BE619424 BE275863 BE394315 BE408109 BE541866 BE253772 BE618236 BE535261
BE296490 BE278212 BE563154 BE257245 BE262274 BE513032 BE378567 BE394152 BE618947 BE269302 BE546516
BE536792 BE615187 BE261186 BE615367 BE619289 BE261184 T49376 AL031671 BE273400 BE563457 BE545597
BE615169 AA150323 AA158723 AA079033 BE313333 AA160100 BE271115 BE294302 BE273051 BE273048 BE622390
AA837947 BE387721 AW973277 AA808731 BE280792 AA160444 BE256723 AI745420 AA643017 BE549441 BE293858
AW975249 AI620819 AW089494 AI434549 BE305231 AA081262 BE280101 AA522507 AI950880 AA187460 BE386860
AW859229 BE170489 BE620149 BE548218 AA316696 AA484426 AI567740 AA160605 AW939805 AA089573 BE300194
BE391331 AW975419 H26808 BE545544 BE615974 AW800241 BE616222 W17343 BE387865 T53697 C03943 BE617637
BE315130 T52942 T50588 N74693 AA187107 T59919 AW797397 AA206447 AA854619 T57175 AI570296 AW517964
AA158269 AI282220 W25297 AI580710 BE262453 AI185868 AA526485 AI288051 AI582513 AA100675 AW615567
BE395354 AI472725 BE314881 BE621281 N99921 AI282689 AI432725 AW732011 AA872254 BE205807 T59435 AI282712
AA650505 AI004374 AA725260 BE313161 T60173 AI371260 BE385641 AW751812 AA078827 AI491858 AI433622
AA219118 AI002092 AA996003 AA064604 AI250287 AI304397 AI453213 AA653630 AI524573 AI440306 H48802 AA157843
AA715629 AW973788 AA932493 AI347563 AA181309 T67880 AA643033 AW467498 AA115904 AA935410 AA483032
AA084568 W25246 AI567588 AA155732 AA158614 AA888319 AA158568 AA188422 AI309183 AA084817 AA157995
AI859659 AA188008 AI287379 AI540675 AA085212 AW028391 AA173297 BE256792 AA182854 BE378771 BE538571
AA079037 BE281597 AA643926 W61011 AA159344 AA320691 AA877597 T57107 AW263819 AI690413 AI619605 AI687579
AA970560 AI368942 AI927104 AW419220 AI620051 AA128490 AA120825 AA079520 AA199648 AW188403 BE045224
AW265533 AA074338 AA102685 AW779399 AA192451 AA182771 AW366812 BE281418 AA211094 AA131073 AA487924
AW674848 AI568103 AA171934 F30349 AW088785 AA581370 AA205482 AW352296 AW517565 AI376249 AA158884
AI340509 T59965 AA085193 AA071570 AI874045 AA852755 BE045217 AW189428 AA211141 AA652134 AI497729
AA994817 AI811459 BE535857 AW769697 AW167892 AW149305 AI864981 AW272126 AW023245 AI439266 AI953196
AA160912 AI718580 BE537547 AA501448 AA069308 L07393 AA353007 AA079235 AI539140 AA740154 W58341 AA888403
BE299000 AA196413 BE613327 BE261523 AA866599 AW844713 AI691159 AI079975 AW327479 BE180731 AA984805
AW500732 AW504061
AA774672 AW504164
303701 1155179_1
303759 447287_1
303773 356632_1
303778 174437_1
303784 414659_1
303845 50211_2
303898 162688_3
AA769074 AA570769 AA808585 AA808682
AW505368 AA218610 F11852 T65345 AA397806
BE297711 AW505574 AA704983
F07942 T08033
BE386266 BE148823 T23215 AI906290 AA299906 BE207197 AW074114 AI760368 AI005358 AW662201 AA188988
AI690711 AA775103 AW072931 AI684269 AW129364 AW615634 AI049941 AW874040 AI352633 AA188989 AI287775
AA868774 AA599660
AA780365 AA909233 AI275542
AA210878 AA215684 R11101
M13560 AA336951 AA161015 R72814 T69687 R75705 T61319 AA158454 R50579 T56649 AI214156 T70375 R31655
H64997 AW800487 H49110 AA634206 H42384 H21783 AI560152 AA664230 H42302 R48708 AA013277 T61901 T92417
AA875985 T61962 T63055 AA430725 AA458964 AA578746 AI582385 T63000 AI499875 H64998 AA022538 AI364804
AI865211 AI439714 AI224059 AI249917 T59258 AA477806 AA715834 AA916120 R38304 R35899 R82985 H25524 H82984
AW516728 T54642 AA079866 H27555 AA455820 T63919 R79450 AI431241 AA937349 AA127213 AA421729 H61196
T63894 AA013050 AA079133 W96364 AA487926 AI762796 H26377 AI433386 AI865423 AW371475 R98189 AA643978
AI718204 AW381954 AI862735
AA323758 R12731 R14082
R17531 AW960899 AA338366 AW673294 BE047729 BE047722 AA330746 AW841797 H05030 AI142105 R12654
H07989 AJ239462 H24544 AA078369 R74153
BE512926 BE304794 AA129140 AA052922 AA092258 BE378058 BE615391 BE615218 BE616188 AI214126 H05675
W56857 AI028525 BE617241 BE531271 AW856227 T56489 AA322005 AW794148 AF170577 BE615738 AA005138 L76930
L76932 L76933 X95410 AW389462 BE563092 AW997937 AA263158 AI520992 AW947350 AA522535 AW945921 AV653776
AW884835 AW947338 AI687178 AW945799 AI905627 AW948449 AV653751 AW945924 AA563898 AW945810 AW945832
AW371449 AW945864 AW948447 AW945910 AA643002 AA522680 AA522715 AA578840 AA523279 AA826150 AW945809
AW405998 AA551909 R23173 AA595545 AW389497 AI933770 AI125053 AI471803 AW795856 AW796937 W30675 H70317

		H68296 T59240 AA397650 H59852 AA938072 AA978010 R35643 T89735 AW361585 AW196153 AI538069 AA604540 AI434259 R49181 T58717 AW062486 AW796966 AI648384 R77733 AI623502 BE171342 BE171303 R35658 AW974883 AW149898 AI500045 AI540710 AI540392 AW009172 AW277199 AI371312 AI500096 AI470297 AW372940 AW844562 AW844560 AW797965 AI691146 X07062 AW799199 H60666 AA837684 AF130734 T25952 AI933771 AI914860 AW391925 AW793843 AW795012 AW366709 AW750987 AW750985 R35765 AW844942 AW750986 H64920 R34651 X86703 5 321039 26338_2 BE018103 BE018083 BE293253 AW247083 BE207643 BE514793 BE183238 AA376427 AW273850 AW043786 BE439973 AL045428 AI889050 AA026496 AI422924 AI884485 W96068 AA020872 F37119 AA714378 AA021107 AA011141 AI554001 AI375841 AI469097 AA335219 AW967315 AI692177 AA410448 AI568858 AA582647 AA026419 AA281639 AW515248 10 AW007777 AA010840 AW188439 AI805423 AI148210 BE301590 AA744414 AA745392 AW167423 AA622659 AW000878 AI432387 AA760930 BE047189 AA021605 AV658045 AI093347 AA588594 H63143 AA639556 AI308976 AA379270 AA633407 AI874329 AI206484 AI493895 AI694103 AI249682 AA973765 AA872445 AI125446 AA287272 AW069761 AA682569 AW009712 BE542774 R50167 BE301574 AA991202 AA502006 AI219819 AW074373 AA617996 AI521242 F25241 15 AW615812 R16774 AA335218 AW673800 H26778 AI468557 AI886986 AI560759 AI460075 AA502968 AA503273 AA610680 AA287274 AA554020 AA284889 AA916636 AW469457 AW273250 AW673708 AW512948 AL041071 AI446042 AA903535 BE172441 AI282411 AW265021 AA810799 AI559865 AA729332 AW004611 AW129451 AA659019 BE208239 AA610825 H03511 BE383995 R16474 AA281701 AW009244 AA287424 AA558139 AW364081 306051 19085_3 F08147 AW408359 AW949429 R23785 AW247442 AA305512 T29095 AA905130 BE246361 BE244981 AA220199 BE504058 X80878 AA533727 AA608601 AW005964 AI811627 AI367037 AI277985 AI493719 AI277848 AA854982 AW247298 AI216345 20 AI041295 AI887378 AA781241 AI674270 AW628959 AI383083 BE504391 AA729421 AA552188 AA373387 AW880360 AW875262 AW875369 AW581540 AW875358 AW581568 R23735 AW134768 321163 171122_1 W03912 AW971410 AA506385 AA209530 H73495 H48629 W56149 321235 1102181_1 H56752 AW340384 N49521 320603 4297_1 AA853680 AK001668 BE386425 BE563549 BE296124 BE298950 R51419 U46295 BE147292 AA360056 R48018 AW845348 25 N47383 AI817280 AI671902 AA988104 AA479464 N56996 AI192374 AI927558 AA659888 AI799903 AA548397 AI161167 AI656333 AI418829 AW592671 BE327906 AW513346 AI888579 AW469410 AW512809 D25682 AA576079 AA479354 T30342 R51307 T16044 H29063 AW079357 AI339477 R47914 AI986068 AI870065 AI868489 AI521099 AI582732 AA995540 AW957299 AA352608 AA676752 AA410510 AA358874 AI865724 AA853679 AI699265 AW188789 N47380 AA233715 BE258194 R55421 R55643 H42362 AA243884 30 AW886407 AA489268 R57015 R58094 BE077459 BE077423 BE546995 AW849216 T69383 AW938111 H60337 BE221073 AB033100 AA347036 BE260325 AW961669 AL047207 AA347037 AI766894 AA601045 AI559897 AW139033 AW274622 AW172884 AW089070 AA804340 AW798925 305704 464759_-1 AA825266 322011 23158_1 AL137354 AL043375 306407 AA971985 35 306454 AA977992 306516 AA989542 306518 AA989598 306526 AA989713 306534 AA991487 40 306590 AI000246 306591 AI000248 306631 AI001149 306654 AI003654 306786 AI041589 45 306799 AI051696 308023 AI452732 308070 AI470948 308099 AI475914 306805 AI055966 50 306814 AI066577 306873 AI086929 306911 AI095365 306982 AI127883 308238 AI559492 55 308258 AI565612 308289 AI571211 308311 AI581855 308332 AI591235 308511 AI687580 60 308601 AI719930 308612 AI735634 308636 AI744063 308814 AI819263 308851 AI829820 65 308981 AI873242 310570 1071946_1 AI318327 AI318328 AI318495 305022 AA627416 305060 AA635771 305070 AA639783
--	--	---

	305079	AA641329
	305134	AA653159
	303977	AW512978
5	305216	AA669056
	305263	AA679467
	305266	AA679772
	305396	AA721052
	305403	AA723748
10	305488	AA749000
	305549	AA773530
	305601	AA780975
	305610	AA782319
	305621	AA789095
15	305710	AA826544
	305724	AA827608
	305744	AA831819
	305752	AA835278
	307018	AI140639
	307055	AI148477
20	307058	AI148709
	305801	AA845997
	305830	AA857665
	305836	AA858043
25	305852	AA862455
	305858	AA863103
	305866	AA864533
	305867	AA864572
	307126	AI184951
30	305903	AA873085
	328803	c_7_hs
	328809	c_7_hs
	305949	AA884409
	328829	c_7_hs
35	330021	c16_p2
	330024	c16_p2
	330028	c16_p2
	330049	c17_p2
	305993	AA889197
40	330095	c19_p2
	330096	c19_p2
	307205	AI192479
	307427	AI243437
	307491	AI268539
45	307581	AI284415
	307588	AI285535
	337672	CH22_6002FG_LINK_EM:AC00
	337693	CH22_6030FG_LINK_EM:AC00
	337738	CH22_6083FG_LINK_EM:AC00
50	307692	AI318342
	307806	AI351739
	309107	AI925823
	309230	AI970747
	339338	CH22_8300FG_LINK_BA3541I
55	309257	AI984183
	309366	AW072970
	309422	AW087175
	325207	c10_hs
	325257	c11_hs
60	309646	AW194694
	309651	AW195850
	325313	c11_hs
	309924	AW340812
	334030	CH22_1308FG_320_2_LINK_EM
65	334040	CH22_1318FG_322_8_LINK_EM
	334083	CH22_1361FG_327_38_LINK_E
	332810	CH22_26FG_7_12_LINK_C65E1
	302747	32813_1 AF062275 L03830
	302753	33029_1 M74299 M74302 M74303
	302777	33803_1 AJ230640 AJ230648

	304094	H11295
	302824 35372_1	U21260 U21258
	302996 41196_1	AF054663 AF124197 R70292
5	325870 c16_hs	
	304240	AA009802
	304410	AA284508
	304443	AA399444
	304475	AA428879
10	304522	AA465405
	304678	AA548556
	304705	AA564064
	306004	AA889992
	306008	AA894390
15	306013	AA896990
	306082	AA908508
	336174 CH22_3567FG_710_1_LINK_DA	
	306094	AA908877
	304823	AA584837
20	304872	AA595289
	304918	AA602697
	304955	AA613504
	306249	AA933840
	306286	AA936892
25	306295	AA937331
	306317	AA947909
	306347	AA961144
	306365	AA962086
	306398	AA970548
30	330401 entrez_D28383	D28383
	330463 460_2	NM_001055 AA332948 U26309 U09031 L19955 L10819 AI366043 X84654 U71086 AV654451 AJ007418 AA053625
		BE168856 AA376730 H12694 AA810348 AA621972 AI818950 AV645367 AI819966 AA910602 AW512449 H67893 AI310497
		AI304330 AI339217 AW193588 AW438688 AI818970 AW316799 AA906527 AA777570 N47673 AI336428 AW945133
		AI038606 R29692 AW194197 AI304748 H12639 AA053178 AA493213 AA676958 AA113154 AI313469 AI368239 R93183
		W24532 U52852 U54701 AL046864 AA365795
35	330535 1374_-8	U11872
	332634 10404_2	U24488 NM_007116

TABLE 13B shows the genomic positioning for those primekeys lacking unigene ID's and accession numbers in Table 13. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique number corresponding to an Eos probeset
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers
 Strand: Indicates DNA strand from which exons were predicted.
 Nt_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	Nt_position
332791	Dunham, I. et.al.	Plus	72720-73315
332792	Dunham, I. et.al.	Plus	73381-73768
332810	Dunham, I. et.al.	Plus	304296-304384
332944	Dunham, I. et.al.	Plus	2414825-2414932
332972	Dunham, I. et.al.	Plus	2572152-2572236
333133	Dunham, I. et.al.	Plus	3360058-3360195
333154	Dunham, I. et.al.	Plus	3615887-3616019
333155	Dunham, I. et.al.	Plus	3616832-3617003
333227	Dunham, I. et.al.	Plus	3992866-3992968
333230	Dunham, I. et.al.	Plus	3995507-3996507
333298	Dunham, I. et.al.	Plus	4581537-4581947
333304	Dunham, I. et.al.	Plus	4629943-4630242
333305	Dunham, I. et.al.	Plus	4630388-4630645
333365	Dunham, I. et.al.	Plus	4786883-4787283
333383	Dunham, I. et.al.	Plus	4907179-4907277
333391	Dunham, I. et.al.	Plus	4916697-4916780
333392	Dunham, I. et.al.	Plus	4918294-4918433
333397	Dunham, I. et.al.	Plus	4922466-4922635
333403	Dunham, I. et.al.	Plus	4925140-4925256
333413	Dunham, I. et.al.	Plus	4943824-4943974
333445	Dunham, I. et.al.	Plus	5097827-5097885
333479	Dunham, I. et.al.	Plus	5272855-5272939
333481	Dunham, I. et.al.	Plus	5286358-5286505
333483	Dunham, I. et.al.	Plus	5297945-5298105
333516	Dunham, I. et.al.	Plus	5570204-5570390
333517	Dunham, I. et.al.	Plus	5570729-5570925
333518	Dunham, I. et.al.	Plus	5571761-5572025
333531	Dunham, I. et.al.	Plus	5622622-5622684
333566	Dunham, I. et.al.	Plus	5954226-5954473
333572	Dunham, I. et.al.	Plus	6026896-6027189
333586	Dunham, I. et.al.	Plus	6246834-6247314
333588	Dunham, I. et.al.	Plus	6255445-6255779
333594	Dunham, I. et.al.	Plus	6308990-6309450
333595	Dunham, I. et.al.	Plus	6323103-6323348
333600	Dunham, I. et.al.	Plus	6355629-6355925
333601	Dunham, I. et.al.	Plus	6360075-6360442
333607	Dunham, I. et.al.	Plus	6504431-6504690
333612	Dunham, I. et.al.	Plus	6549563-6549697
333613	Dunham, I. et.al.	Plus	6550643-6550748
333614	Dunham, I. et.al.	Plus	6551227-6551389
333624	Dunham, I. et.al.	Plus	6595146-6595244
333626	Dunham, I. et.al.	Plus	6614174-6614467
333635	Dunham, I. et.al.	Plus	6663683-6663973
333637	Dunham, I. et.al.	Plus	6674968-6675134
333642	Dunham, I. et.al.	Plus	6708760-6709139
333647	Dunham, I. et.al.	Plus	6772502-6772779
333653	Dunham, I. et.al.	Plus	6811130-6811392
333654	Dunham, I. et.al.	Plus	6816731-6816993
333656	Dunham, I. et.al.	Plus	6822087-6822406
333657	Dunham, I. et.al.	Plus	6831369-6831445
333658	Dunham, I. et.al.	Plus	6835282-6835474

	333659	Dunham, I. et.al.	Plus	6836179-6836248
	333684	Dunham, I. et.al.	Plus	7169561-7169742
	333686	Dunham, I. et.al.	Plus	7177117-7177302
5	333697	Dunham, I. et.al.	Plus	7203859-7203934
	333698	Dunham, I. et.al.	Plus	7205279-7205383
	333699	Dunham, I. et.al.	Plus	7206101-7206175
	333703	Dunham, I. et.al.	Plus	7215559-7215663
	333709	Dunham, I. et.al.	Plus	7229730-7229835
	333747	Dunham, I. et.al.	Plus	7605884-7606206
10	333774	Dunham, I. et.al.	Plus	7716509-7716636
	333775	Dunham, I. et.al.	Plus	7729983-7730149
	333806	Dunham, I. et.al.	Plus	7877475-7877666
	333843	Dunham, I. et.al.	Plus	7978762-7978887
15	333854	Dunham, I. et.al.	Plus	8029446-8029524
	333873	Dunham, I. et.al.	Plus	8133266-8133429
	333880	Dunham, I. et.al.	Plus	8151923-8152133
	333885	Dunham, I. et.al.	Plus	8154352-8154437
	333918	Dunham, I. et.al.	Plus	8307124-8307215
20	333947	Dunham, I. et.al.	Plus	8579888-8579966
	333961	Dunham, I. et.al.	Plus	8617999-8618104
	333981	Dunham, I. et.al.	Plus	8782374-8782643
	333991	Dunham, I. et.al.	Plus	8837419-8837551
	333994	Dunham, I. et.al.	Plus	8852749-8852894
25	334030	Dunham, I. et.al.	Plus	9288463-9288782
	334083	Dunham, I. et.al.	Plus	9837016-9837081
	334111	Dunham, I. et.al.	Plus	10279365-10279531
	334135	Dunham, I. et.al.	Plus	10457085-10457183
	334218	Dunham, I. et.al.	Plus	12680289-12680378
30	334249	Dunham, I. et.al.	Plus	13190430-13190574
	334262	Dunham, I. et.al.	Plus	13231452-13231581
	334264	Dunham, I. et.al.	Plus	13234447-13234544
	334327	Dunham, I. et.al.	Plus	13577413-13577496
	334328	Dunham, I. et.al.	Plus	13589868-13589936
	334340	Dunham, I. et.al.	Plus	13642407-13642522
35	334454	Dunham, I. et.al.	Plus	14326506-14326738
	334504	Dunham, I. et.al.	Plus	14510206-14510398
	334508	Dunham, I. et.al.	Plus	14514936-14515122
	334512	Dunham, I. et.al.	Plus	14545933-14546366
40	334582	Dunham, I. et.al.	Plus	15026255-15026371
	334659	Dunham, I. et.al.	Plus	15460624-15460726
	334721	Dunham, I. et.al.	Plus	15796816-15796987
	334723	Dunham, I. et.al.	Plus	15805317-15805399
	334730	Dunham, I. et.al.	Plus	15967830-15967934
	334774	Dunham, I. et.al.	Plus	16251857-16252178
45	334778	Dunham, I. et.al.	Plus	16276180-16276395
	334851	Dunham, I. et.al.	Plus	17820110-17820810
	334885	Dunham, I. et.al.	Plus	19233667-19233787
	334902	Dunham, I. et.al.	Plus	19317083-19317195
	334905	Dunham, I. et.al.	Plus	19322553-19322680
50	334906	Dunham, I. et.al.	Plus	19323493-19323590
	334910	Dunham, I. et.al.	Plus	19398155-19398684
	335018	Dunham, I. et.al.	Plus	20688288-20688415
	335025	Dunham, I. et.al.	Plus	20743941-20744050
55	335033	Dunham, I. et.al.	Plus	20753188-20753314
	335044	Dunham, I. et.al.	Plus	20842088-20842682
	335142	Dunham, I. et.al.	Plus	21465105-21465186
	335157	Dunham, I. et.al.	Plus	21543302-21544341
	335160	Dunham, I. et.al.	Plus	21573388-21573497
60	335174	Dunham, I. et.al.	Plus	21631301-21631447
	335188	Dunham, I. et.al.	Plus	21669118-21669328
	335190	Dunham, I. et.al.	Plus	21680807-21680876
	335191	Dunham, I. et.al.	Plus	21681110-21681183
	335193	Dunham, I. et.al.	Plus	21692208-21692362
65	335204	Dunham, I. et.al.	Plus	21750636-21750726
	335222	Dunham, I. et.al.	Plus	21885542-21885608
	335226	Dunham, I. et.al.	Plus	21890838-21890930
	335227	Dunham, I. et.al.	Plus	21892145-21892289
	335309	Dunham, I. et.al.	Plus	22500158-22500276
	335310	Dunham, I. et.al.	Plus	22500714-22500831

	335311	Dunham, I. et.al.	Plus	22501602-22501676
	335355	Dunham, I. et.al.	Plus	22779222-22779516
	335362	Dunham, I. et.al.	Plus	22809167-22809461
5	335368	Dunham, I. et.al.	Plus	22843040-22843184
	335384	Dunham, I. et.al.	Plus	22918150-22918263
	335385	Dunham, I. et.al.	Plus	22919072-22919339
	335436	Dunham, I. et.al.	Plus	23427793-23427923
	335440	Dunham, I. et.al.	Plus	23458702-23459017
10	335441	Dunham, I. et.al.	Plus	23460632-23460724
	335450	Dunham, I. et.al.	Plus	23480190-23480270
	335453	Dunham, I. et.al.	Plus	23483333-23483459
	335458	Dunham, I. et.al.	Plus	23490034-23490143
	335464	Dunham, I. et.al.	Plus	23500331-23500496
15	335496	Dunham, I. et.al.	Plus	24164386-24164545
	335497	Dunham, I. et.al.	Plus	24167666-24167869
	335498	Dunham, I. et.al.	Plus	24172082-24172161
	335499	Dunham, I. et.al.	Plus	24176698-24176869
	335500	Dunham, I. et.al.	Plus	24178236-24178326
20	335507	Dunham, I. et.al.	Plus	24219973-24220039
	335510	Dunham, I. et.al.	Plus	24222975-24223118
	335513	Dunham, I. et.al.	Plus	24224272-24224496
	335627	Dunham, I. et.al.	Plus	25150005-25150061
	335651	Dunham, I. et.al.	Plus	25317560-25317696
25	335655	Dunham, I. et.al.	Plus	25333211-25333369
	335656	Dunham, I. et.al.	Plus	25333601-25333751
	335658	Dunham, I. et.al.	Plus	25336315-25336406
	335663	Dunham, I. et.al.	Plus	25342680-25342802
	335665	Dunham, I. et.al.	Plus	25344096-25344287
30	335667	Dunham, I. et.al.	Plus	25345735-25345856
	335668	Dunham, I. et.al.	Plus	25346313-25346447
	335689	Dunham, I. et.al.	Plus	25454350-25454604
	335690	Dunham, I. et.al.	Plus	25455442-25455625
	335715	Dunham, I. et.al.	Plus	25565941-25566052
35	335719	Dunham, I. et.al.	Plus	25593936-25594101
	335734	Dunham, I. et.al.	Plus	25688723-25688869
	335744	Dunham, I. et.al.	Plus	25716483-25716615
	335809	Dunham, I. et.al.	Plus	26310772-26310909
	335819	Dunham, I. et.al.	Plus	26356341-26356470
40	335822	Dunham, I. et.al.	Plus	26364087-26364196
	335872	Dunham, I. et.al.	Plus	26820760-26820943
	335885	Dunham, I. et.al.	Plus	26933436-26933534
	335968	Dunham, I. et.al.	Plus	27743843-27744029
	335971	Dunham, I. et.al.	Plus	27752808-27753017
45	335975	Dunham, I. et.al.	Plus	27801321-27801391
	335976	Dunham, I. et.al.	Plus	27809041-27809187
	335989	Dunham, I. et.al.	Plus	27983788-27983860
	335990	Dunham, I. et.al.	Plus	27988532-27988608
	336010	Dunham, I. et.al.	Plus	28570239-28570330
50	336093	Dunham, I. et.al.	Plus	29556922-29557002
	336126	Dunham, I. et.al.	Plus	30057891-30058105
	336129	Dunham, I. et.al.	Plus	30062259-30062348
	336187	Dunham, I. et.al.	Plus	30433494-30433585
	336188	Dunham, I. et.al.	Plus	30434870-30435004
55	336225	Dunham, I. et.al.	Plus	30833614-30833788
	336371	Dunham, I. et.al.	Plus	33968108-33968204
	336373	Dunham, I. et.al.	Plus	33976308-33976504
	336377	Dunham, I. et.al.	Plus	33994489-33994599
	336380	Dunham, I. et.al.	Plus	33995323-33995434
60	336383	Dunham, I. et.al.	Plus	34005784-34005964
	336384	Dunham, I. et.al.	Plus	34007429-34007559
	336385	Dunham, I. et.al.	Plus	34007879-34008159
	336386	Dunham, I. et.al.	Plus	34012965-34013115
	336441	Dunham, I. et.al.	Plus	34187606-34187663
	336444	Dunham, I. et.al.	Plus	34190585-34190718
65	336484	Dunham, I. et.al.	Plus	34237425-34237505
	336497	Dunham, I. et.al.	Plus	34267190-34267245
	336499	Dunham, I. et.al.	Plus	34267504-34267572
	336503	Dunham, I. et.al.	Plus	34271306-34271372
	336548	Dunham, I. et.al.	Plus	34353881-34354826

	336552	Dunham, I. et.al.	Plus	34356420-34356527
	336553	Dunham, I. et.al.	Plus	34356683-34356753
	336567	Dunham, I. et.al.	Plus	34428228-34428395
5	336568	Dunham, I. et.al.	Plus	34428521-34428637
	336659	Dunham, I. et.al.	Plus	1896402-1896478
	336715	Dunham, I. et.al.	Plus	3110198-3110314
	336803	Dunham, I. et.al.	Plus	6106904-6106990
	336805	Dunham, I. et.al.	Plus	6126661-6126786
10	336850	Dunham, I. et.al.	Plus	7745284-7745355
	336857	Dunham, I. et.al.	Plus	8130457-8130612
	336911	Dunham, I. et.al.	Plus	11035818-11035984
	336949	Dunham, I. et.al.	Plus	12818687-12818891
	336950	Dunham, I. et.al.	Plus	12875843-12875912
15	336958	Dunham, I. et.al.	Plus	13203550-13203973
	336993	Dunham, I. et.al.	Plus	15096270-15096324
	337076	Dunham, I. et.al.	Plus	19338177-19338679
	337109	Dunham, I. et.al.	Plus	21166580-21166650
	337123	Dunham, I. et.al.	Plus	22052874-22052942
20	337151	Dunham, I. et.al.	Plus	23106433-23106510
	337189	Dunham, I. et.al.	Plus	24225887-24225954
	337241	Dunham, I. et.al.	Plus	27280182-27280313
	337337	Dunham, I. et.al.	Plus	30395182-30395285
	337353	Dunham, I. et.al.	Plus	30804624-30804780
25	337384	Dunham, I. et.al.	Plus	31333399-31333580
	337396	Dunham, I. et.al.	Plus	31585902-31586067
	337414	Dunham, I. et.al.	Plus	31953012-31953205
	337418	Dunham, I. et.al.	Plus	32014049-32014131
	337461	Dunham, I. et.al.	Plus	32803968-32804028
30	337480	Dunham, I. et.al.	Plus	33219714-33219779
	337482	Dunham, I. et.al.	Plus	33227865-33227946
	337483	Dunham, I. et.al.	Plus	33237292-33237427
	337490	Dunham, I. et.al.	Plus	33318571-33318644
	337522	Dunham, I. et.al.	Plus	33963188-33963979
35	337532	Dunham, I. et.al.	Plus	34187269-34187366
	337552	Dunham, I. et.al.	Plus	19497-19600
	337584	Dunham, I. et.al.	Plus	945236-945452
	337611	Dunham, I. et.al.	Plus	1482883-1483016
40	337672	Dunham, I. et.al.	Plus	3331236-3331313
	337693	Dunham, I. et.al.	Plus	3575975-3576153
	337738	Dunham, I. et.al.	Plus	3865738-3865814
	337926	Dunham, I. et.al.	Plus	6286377-6286470
	337927	Dunham, I. et.al.	Plus	6343033-6343172
	337935	Dunham, I. et.al.	Plus	6534661-6534782
45	337944	Dunham, I. et.al.	Plus	6589383-6589450
	337954	Dunham, I. et.al.	Plus	6831483-6831620
	337996	Dunham, I. et.al.	Plus	7445532-7445633
	338004	Dunham, I. et.al.	Plus	7601363-7601520
	338016	Dunham, I. et.al.	Plus	7863131-7863310
50	338174	Dunham, I. et.al.	Plus	12771102-12771268
	338176	Dunham, I. et.al.	Plus	12774072-12774223
	338238	Dunham, I. et.al.	Plus	14661936-14662015
	338277	Dunham, I. et.al.	Plus	16167622-16167962
	338294	Dunham, I. et.al.	Plus	16463958-16464539
55	338316	Dunham, I. et.al.	Plus	17089711-17089988
	338323	Dunham, I. et.al.	Plus	17154655-17154792
	338324	Dunham, I. et.al.	Plus	17155309-17155574
	338386	Dunham, I. et.al.	Plus	18611213-18611407
	338398	Dunham, I. et.al.	Plus	18953492-18953581
60	338410	Dunham, I. et.al.	Plus	19292807-19292916
	338414	Dunham, I. et.al.	Plus	19345573-19345660
	338460	Dunham, I. et.al.	Plus	20233372-20233488
	338481	Dunham, I. et.al.	Plus	20942659-20942873
	338489	Dunham, I. et.al.	Plus	21142605-21143049
65	338500	Dunham, I. et.al.	Plus	21253847-21253974
	338514	Dunham, I. et.al.	Plus	21379420-21379655
	338530	Dunham, I. et.al.	Plus	21636361-21636509
	338620	Dunham, I. et.al.	Plus	23540239-23540334
	338631	Dunham, I. et.al.	Plus	23711167-23711241
	338653	Dunham, I. et.al.	Plus	24219427-24219509

	338660	Dunham, I. et.al.	Plus	24387122-24387266
	338704	Dunham, I. et.al.	Plus	25230432-25230548
	338847	Dunham, I. et.al.	Plus	27995337-27995420
5	338887	Dunham, I. et.al.	Plus	28465244-28465384
	338895	Dunham, I. et.al.	Plus	28598893-28599135
	338915	Dunham, I. et.al.	Plus	28824881-28824977
	338925	Dunham, I. et.al.	Plus	28883892-28884036
	338936	Dunham, I. et.al.	Plus	29148022-29148160
10	338952	Dunham, I. et.al.	Plus	29418831-29418968
	338980	Dunham, I. et.al.	Plus	29896789-29896874
	338981	Dunham, I. et.al.	Plus	29897917-29898008
	338986	Dunham, I. et.al.	Plus	30007287-30007415
	339009	Dunham, I. et.al.	Plus	30348477-30348598
15	339017	Dunham, I. et.al.	Plus	30420896-30421090
	339045	Dunham, I. et.al.	Plus	30744286-30744356
	339046	Dunham, I. et.al.	Plus	30746269-30746420
	339059	Dunham, I. et.al.	Plus	30814655-30814801
	339067	Dunham, I. et.al.	Plus	30869347-30869412
20	339069	Dunham, I. et.al.	Plus	30880975-30881070
	339078	Dunham, I. et.al.	Plus	30914310-30914423
	339084	Dunham, I. et.al.	Plus	30944556-30944803
	339101	Dunham, I. et.al.	Plus	31158047-31158123
	339102	Dunham, I. et.al.	Plus	31169321-31169563
	339103	Dunham, I. et.al.	Plus	31170343-31170454
25	339115	Dunham, I. et.al.	Plus	31459869-31459927
	339157	Dunham, I. et.al.	Plus	32131701-32131833
	339166	Dunham, I. et.al.	Plus	32210902-32211006
	339167	Dunham, I. et.al.	Plus	32213567-32213730
30	339288	Dunham, I. et.al.	Plus	33169611-33169691
	339289	Dunham, I. et.al.	Plus	33186756-33186903
	339291	Dunham, I. et.al.	Plus	33205057-33205247
	339407	Dunham, I. et.al.	Plus	34189461-34189620
	332865	Dunham, I. et.al.	Minus	1391482-1391218
35	332881	Dunham, I. et.al.	Minus	1563520-1563184
	332930	Dunham, I. et.al.	Minus	2022565-2022497
	332931	Dunham, I. et.al.	Minus	2023651-2023562
	332984	Dunham, I. et.al.	Minus	2632606-2632457
	332986	Dunham, I. et.al.	Minus	2635398-2635206
40	332997	Dunham, I. et.al.	Minus	2710509-2710375
	333051	Dunham, I. et.al.	Minus	2991973-2991840
	333061	Dunham, I. et.al.	Minus	3029631-3029527
	333064	Dunham, I. et.al.	Minus	3030722-3030623
	333096	Dunham, I. et.al.	Minus	3184234-3184118
45	333099	Dunham, I. et.al.	Minus	3206796-3206674
	333106	Dunham, I. et.al.	Minus	3230744-3230547
	333160	Dunham, I. et.al.	Minus	3654893-3654678
	333163	Dunham, I. et.al.	Minus	3665124-3664962
	333165	Dunham, I. et.al.	Minus	3674052-3673905
50	333166	Dunham, I. et.al.	Minus	3694664-3694567
	333170	Dunham, I. et.al.	Minus	3733394-3733299
	333174	Dunham, I. et.al.	Minus	3764284-3764210
	333188	Dunham, I. et.al.	Minus	3826990-3826863
	333214	Dunham, I. et.al.	Minus	3966559-3966437
55	333232	Dunham, I. et.al.	Minus	4001551-4001365
	333237	Dunham, I. et.al.	Minus	4003326-4003219
	333239	Dunham, I. et.al.	Minus	4095861-4094462
	333255	Dunham, I. et.al.	Minus	4297883-4297716
	333259	Dunham, I. et.al.	Minus	4306769-4306639
	333274	Dunham, I. et.al.	Minus	4389146-4388954
60	333290	Dunham, I. et.al.	Minus	4530734-4530554
	333295	Dunham, I. et.al.	Minus	4549290-4549198
	333296	Dunham, I. et.al.	Minus	4550766-4550644
	333310	Dunham, I. et.al.	Minus	4637315-4637232
	333311	Dunham, I. et.al.	Minus	4637933-4637844
65	333312	Dunham, I. et.al.	Minus	4638794-4638635
	333313	Dunham, I. et.al.	Minus	4639397-4639277
	333315	Dunham, I. et.al.	Minus	5405980-5405876
	333318	Dunham, I. et.al.	Minus	4642636-4642564
	333321	Dunham, I. et.al.	Minus	4649080-4648934

	333327	Dunham, I. et.al.	Minus	4657947-4657828
	333335	Dunham, I. et.al.	Minus	4672656-4672564
	333337	Dunham, I. et.al.	Minus	4677930-4677841
5	333454	Dunham, I. et.al.	Minus	5137007-5136880
	333458	Dunham, I. et.al.	Minus	5143942-5143806
	333459	Dunham, I. et.al.	Minus	5144548-5144344
	333470	Dunham, I. et.al.	Minus	5223319-5223088
	333493	Dunham, I. et.al.	Minus	4637315-4637232
10	333496	Dunham, I. et.al.	Minus	5404643-5404523
	333498	Dunham, I. et.al.	Minus	5405980-5405876
	333510	Dunham, I. et.al.	Minus	5557628-5557469
	333546	Dunham, I. et.al.	Minus	5886643-5886442
	333561	Dunham, I. et.al.	Minus	5903659-5903590
15	333738	Dunham, I. et.al.	Minus	7552160-7552084
	333780	Dunham, I. et.al.	Minus	7750367-7750277
	333783	Dunham, I. et.al.	Minus	7751850-7751777
	333818	Dunham, I. et.al.	Minus	7911959-7911762
	333894	Dunham, I. et.al.	Minus	8188855-8188709
20	333897	Dunham, I. et.al.	Minus	8194390-8194284
	333900	Dunham, I. et.al.	Minus	8200268-8200122
	333909	Dunham, I. et.al.	Minus	8229639-8229477
	333936	Dunham, I. et.al.	Minus	8512805-8512564
	333944	Dunham, I. et.al.	Minus	8557051-8556936
25	334040	Dunham, I. et.al.	Minus	9342995-9342934
	334154	Dunham, I. et.al.	Minus	10570714-10570572
	334178	Dunham, I. et.al.	Minus	11755052-11754971
	334188	Dunham, I. et.al.	Minus	11925963-11925834
	334273	Dunham, I. et.al.	Minus	13265608-13265522
30	334282	Dunham, I. et.al.	Minus	13285293-13285178
	334285	Dunham, I. et.al.	Minus	13289990-13289793
	334286	Dunham, I. et.al.	Minus	13291759-13291569
	334303	Dunham, I. et.al.	Minus	13454331-13454217
	334305	Dunham, I. et.al.	Minus	13456310-13456209
35	334306	Dunham, I. et.al.	Minus	13461157-13461049
	334320	Dunham, I. et.al.	Minus	13496857-13496717
	334352	Dunham, I. et.al.	Minus	13675908-13675828
	334353	Dunham, I. et.al.	Minus	13683722-13683596
	334359	Dunham, I. et.al.	Minus	13728664-13728534
40	334363	Dunham, I. et.al.	Minus	13740004-13739812
	334365	Dunham, I. et.al.	Minus	13742078-13741971
	334399	Dunham, I. et.al.	Minus	14186289-14186163
	334409	Dunham, I. et.al.	Minus	14195181-14195075
	334414	Dunham, I. et.al.	Minus	14234033-14233932
45	334470	Dunham, I. et.al.	Minus	14389581-14389442
	334483	Dunham, I. et.al.	Minus	14428355-14428281
	334489	Dunham, I. et.al.	Minus	14455428-14454288
	334498	Dunham, I. et.al.	Minus	14483789-14483700
	334501	Dunham, I. et.al.	Minus	14487509-14487356
50	334502	Dunham, I. et.al.	Minus	14488605-14488526
	334543	Dunham, I. et.al.	Minus	14834496-14834116
	334622	Dunham, I. et.al.	Minus	15191678-15191609
	334650	Dunham, I. et.al.	Minus	15371251-15371178
	334680	Dunham, I. et.al.	Minus	15520047-15519887
55	334745	Dunham, I. et.al.	Minus	16049960-16049653
	334756	Dunham, I. et.al.	Minus	16128678-16128528
	334758	Dunham, I. et.al.	Minus	16132368-16132233
	334761	Dunham, I. et.al.	Minus	16138424-16138319
	334763	Dunham, I. et.al.	Minus	16148136-16148077
60	334784	Dunham, I. et.al.	Minus	16294548-16294360
	334790	Dunham, I. et.al.	Minus	16307576-16307509
	334793	Dunham, I. et.al.	Minus	16330748-16330681
	334802	Dunham, I. et.al.	Minus	16413158-16413026
	334820	Dunham, I. et.al.	Minus	16764338-16764249
65	334824	Dunham, I. et.al.	Minus	16857777-16857674
	334832	Dunham, I. et.al.	Minus	17173957-17173760
	334842	Dunham, I. et.al.	Minus	17464352-17464181
	334844	Dunham, I. et.al.	Minus	17503891-17503768
	334857	Dunham, I. et.al.	Minus	18488368-18488242
	334927	Dunham, I. et.al.	Minus	19988711-19987853

	334939	Dunham, I. et.al.	Minus	20131162-20131054
	334951	Dunham, I. et.al.	Minus	20147708-20147502
	334969	Dunham, I. et.al.	Minus	20188176-20188020
5	334972	Dunham, I. et.al.	Minus	20294734-20294611
	335050	Dunham, I. et.al.	Minus	20884109-20883951
	335078	Dunham, I. et.al.	Minus	21059529-21059458
	335102	Dunham, I. et.al.	Minus	21313841-21313598
	335105	Dunham, I. et.al.	Minus	21320563-21320440
10	335110	Dunham, I. et.al.	Minus	21334136-21333811
	335111	Dunham, I. et.al.	Minus	21335946-21335809
	335115	Dunham, I. et.al.	Minus	21388250-21388146
	335116	Dunham, I. et.al.	Minus	21388573-21388414
	335185	Dunham, I. et.al.	Minus	21651593-21651522
15	335186	Dunham, I. et.al.	Minus	21656436-21656338
	335230	Dunham, I. et.al.	Minus	21899517-21898678
	335236	Dunham, I. et.al.	Minus	21915016-21914870
	335243	Dunham, I. et.al.	Minus	21933519-21933365
	335249	Dunham, I. et.al.	Minus	21950851-21950669
20	335258	Dunham, I. et.al.	Minus	22043431-22043262
	335261	Dunham, I. et.al.	Minus	22063937-22063772
	335276	Dunham, I. et.al.	Minus	22154036-22153937
	335279	Dunham, I. et.al.	Minus	22168834-22168638
	335330	Dunham, I. et.al.	Minus	22556589-22556422
25	335331	Dunham, I. et.al.	Minus	22556823-22556708
	335334	Dunham, I. et.al.	Minus	22560390-22560136
	335346	Dunham, I. et.al.	Minus	22641097-22640918
	335349	Dunham, I. et.al.	Minus	22661861-22661271
	335611	Dunham, I. et.al.	Minus	25070825-25070706
30	335612	Dunham, I. et.al.	Minus	25072328-25072142
	335671	Dunham, I. et.al.	Minus	25358629-25358533
	335676	Dunham, I. et.al.	Minus	25395274-25395152
	335680	Dunham, I. et.al.	Minus	25402437-25402361
	335750	Dunham, I. et.al.	Minus	25732501-25731972
35	335752	Dunham, I. et.al.	Minus	25757026-25756890
	335755	Dunham, I. et.al.	Minus	25763806-25763747
	335767	Dunham, I. et.al.	Minus	25819547-25819218
	335774	Dunham, I. et.al.	Minus	25883733-25883572
	335777	Dunham, I. et.al.	Minus	25885770-25885599
40	335778	Dunham, I. et.al.	Minus	25886469-25886334
	335797	Dunham, I. et.al.	Minus	25958182-25958030
	335800	Dunham, I. et.al.	Minus	25985373-25985280
	335818	Dunham, I. et.al.	Minus	26323886-26323744
	335834	Dunham, I. et.al.	Minus	26391707-26391530
45	335840	Dunham, I. et.al.	Minus	26420596-26420538
	335844	Dunham, I. et.al.	Minus	26433427-26433344
	335846	Dunham, I. et.al.	Minus	26436727-26436621
	335856	Dunham, I. et.al.	Minus	26662452-26662346
	335887	Dunham, I. et.al.	Minus	26939225-26938782
50	335888	Dunham, I. et.al.	Minus	26943037-26942820
	335889	Dunham, I. et.al.	Minus	26946988-26946901
	335890	Dunham, I. et.al.	Minus	26949087-26948665
	335893	Dunham, I. et.al.	Minus	26973898-26973747
	335895	Dunham, I. et.al.	Minus	26975307-26975239
55	335896	Dunham, I. et.al.	Minus	26977639-26977558
	335900	Dunham, I. et.al.	Minus	26980354-26980238
	335907	Dunham, I. et.al.	Minus	27013352-27013273
	335943	Dunham, I. et.al.	Minus	27446610-27446378
	335956	Dunham, I. et.al.	Minus	27653729-27653635
60	335959	Dunham, I. et.al.	Minus	27682313-27682145
	335962	Dunham, I. et.al.	Minus	27704276-27704144
	336040	Dunham, I. et.al.	Minus	29036458-29036300
	336044	Dunham, I. et.al.	Minus	29043828-29043727
	336047	Dunham, I. et.al.	Minus	29050617-29050466
65	336068	Dunham, I. et.al.	Minus	29252077-29251969
	336143	Dunham, I. et.al.	Minus	30135948-30135854
	336158	Dunham, I. et.al.	Minus	30163730-30163610
	336174	Dunham, I. et.al.	Minus	30241988-30241839
	336223	Dunham, I. et.al.	Minus	30816306-30816195
	336245	Dunham, I. et.al.	Minus	31420569-31420509

	336274	Dunham, I. et.al.	Minus	32085468-32085303
	336318	Dunham, I. et.al.	Minus	33364452-33364338
	336326	Dunham, I. et.al.	Minus	33567328-33567201
5	336339	Dunham, I. et.al.	Minus	33798479-33798330
	336340	Dunham, I. et.al.	Minus	33812069-33811915
	336355	Dunham, I. et.al.	Minus	33874750-33874649
	336392	Dunham, I. et.al.	Minus	34015868-34015736
	336393	Dunham, I. et.al.	Minus	34016145-34015951
10	336394	Dunham, I. et.al.	Minus	34016457-34016298
	336400	Dunham, I. et.al.	Minus	34023437-34023298
	336402	Dunham, I. et.al.	Minus	34024090-34023981
	336413	Dunham, I. et.al.	Minus	34046702-34046576
	336424	Dunham, I. et.al.	Minus	34055549-34055491
15	336425	Dunham, I. et.al.	Minus	34058544-34058446
	336437	Dunham, I. et.al.	Minus	34074154-34074090
	336447	Dunham, I. et.al.	Minus	34198207-34197996
	336449	Dunham, I. et.al.	Minus	34204707-34204577
	336466	Dunham, I. et.al.	Minus	34213195-34213046
20	336492	Dunham, I. et.al.	Minus	34255578-34255437
	336511	Dunham, I. et.al.	Minus	34277480-34277351
	336512	Dunham, I. et.al.	Minus	34278373-34278275
	336520	Dunham, I. et.al.	Minus	34319184-34319101
	336522	Dunham, I. et.al.	Minus	34320169-34320056
25	336524	Dunham, I. et.al.	Minus	34321055-34320921
	336527	Dunham, I. et.al.	Minus	34322071-34321966
	336534	Dunham, I. et.al.	Minus	34326797-34326620
	336536	Dunham, I. et.al.	Minus	34327678-34327538
	336542	Dunham, I. et.al.	Minus	34331316-34331183
30	336556	Dunham, I. et.al.	Minus	34375244-34374907
	336557	Dunham, I. et.al.	Minus	34375443-34375341
	336558	Dunham, I. et.al.	Minus	34375825-34375698
	336559	Dunham, I. et.al.	Minus	34376430-34376261
	336560	Dunham, I. et.al.	Minus	34376814-34376596
35	336561	Dunham, I. et.al.	Minus	34377168-34376928
	336597	Dunham, I. et.al.	Minus	7627912-7627757
	336601	Dunham, I. et.al.	Minus	13265853-13265654
	336642	Dunham, I. et.al.	Minus	1304281-1304212
	336645	Dunham, I. et.al.	Minus	1351268-1351168
40	336662	Dunham, I. et.al.	Minus	2158060-2157993
	336664	Dunham, I. et.al.	Minus	1993558-1993481
	336676	Dunham, I. et.al.	Minus	2022565-2022497
	336684	Dunham, I. et.al.	Minus	2158060-2157993
	336686	Dunham, I. et.al.	Minus	2160698-2160486
45	336714	Dunham, I. et.al.	Minus	3094026-3093871
	336719	Dunham, I. et.al.	Minus	3331631-3331503
	336736	Dunham, I. et.al.	Minus	4093128-4093041
	336744	Dunham, I. et.al.	Minus	4333001-4332848
	336786	Dunham, I. et.al.	Minus	5419973-5419873
50	336793	Dunham, I. et.al.	Minus	5631345-5631237
	336859	Dunham, I. et.al.	Minus	8201756-8201561
	336863	Dunham, I. et.al.	Minus	8396673-8396425
	336933	Dunham, I. et.al.	Minus	11760045-11759981
	336942	Dunham, I. et.al.	Minus	12027537-12027455
55	336960	Dunham, I. et.al.	Minus	13267243-13267172
	336969	Dunham, I. et.al.	Minus	13725722-13725643
	336971	Dunham, I. et.al.	Minus	13732308-13732221
	337003	Dunham, I. et.al.	Minus	15523541-15523422
	337011	Dunham, I. et.al.	Minus	16106423-16106080
60	337070	Dunham, I. et.al.	Minus	19034423-19034321
	337072	Dunham, I. et.al.	Minus	19077452-19077323
	337086	Dunham, I. et.al.	Minus	19657011-19656881
	337140	Dunham, I. et.al.	Minus	22649450-22649388
	337193	Dunham, I. et.al.	Minus	24594969-24594874
	337256	Dunham, I. et.al.	Minus	27659956-27659876
65	337278	Dunham, I. et.al.	Minus	28429017-28428848
	337284	Dunham, I. et.al.	Minus	28491414-28491094
	337293	Dunham, I. et.al.	Minus	28846334-28845873
	337316	Dunham, I. et.al.	Minus	29657129-29656997
	337326	Dunham, I. et.al.	Minus	30017199-30017069

	337382	Dunham, I. et.al.	Minus	31233666-31233579
	337392	Dunham, I. et.al.	Minus	31442311-31442229
	337406	Dunham, I. et.al.	Minus	31864840-31864588
5	337412	Dunham, I. et.al.	Minus	31916487-31916312
	337419	Dunham, I. et.al.	Minus	32021496-32021170
	337436	Dunham, I. et.al.	Minus	32257869-32257739
	337455	Dunham, I. et.al.	Minus	32434517-32434425
	337509	Dunham, I. et.al.	Minus	33414613-33414498
10	337518	Dunham, I. et.al.	Minus	33796750-33796647
	337529	Dunham, I. et.al.	Minus	34043668-34043546
	337533	Dunham, I. et.al.	Minus	34193388-34193261
	337539	Dunham, I. et.al.	Minus	34254490-34254322
	337551	Dunham, I. et.al.	Minus	34524446-34524362
15	337553	Dunham, I. et.al.	Minus	24230-24160
	337591	Dunham, I. et.al.	Minus	1006414-1006184
	337592	Dunham, I. et.al.	Minus	1007791-1007634
	337593	Dunham, I. et.al.	Minus	1009460-1009291
	337607	Dunham, I. et.al.	Minus	1355719-1355637
20	337612	Dunham, I. et.al.	Minus	1570235-1570142
	337635	Dunham, I. et.al.	Minus	2169690-2169569
	337824	Dunham, I. et.al.	Minus	4559540-4559266
	337825	Dunham, I. et.al.	Minus	4567155-4567005
	337850	Dunham, I. et.al.	Minus	5077143-5076943
25	337854	Dunham, I. et.al.	Minus	5153435-5153272
	337913	Dunham, I. et.al.	Minus	6149843-6149786
	337915	Dunham, I. et.al.	Minus	5922748-5922690
	337968	Dunham, I. et.al.	Minus	7095797-7095680
	338010	Dunham, I. et.al.	Minus	7754282-7754184
30	338012	Dunham, I. et.al.	Minus	7761421-7761351
	338017	Dunham, I. et.al.	Minus	7864521-7864401
	338065	Dunham, I. et.al.	Minus	7235048-7234950
	338094	Dunham, I. et.al.	Minus	9595602-9595440
	338129	Dunham, I. et.al.	Minus	10915338-10915237
35	338132	Dunham, I. et.al.	Minus	10989617-10989530
	338150	Dunham, I. et.al.	Minus	11478551-11478355
	338157	Dunham, I. et.al.	Minus	11731444-11731375
	338195	Dunham, I. et.al.	Minus	13484103-13483972
	338255	Dunham, I. et.al.	Minus	15242294-15242231
40	338276	Dunham, I. et.al.	Minus	16109555-16109398
	338431	Dunham, I. et.al.	Minus	19747608-19747496
	338448	Dunham, I. et.al.	Minus	20151152-20151054
	338451	Dunham, I. et.al.	Minus	20174286-20174193
	338477	Dunham, I. et.al.	Minus	20821897-20821838
45	338534	Dunham, I. et.al.	Minus	21771238-21771170
	338682	Dunham, I. et.al.	Minus	24800712-24800461
	338684	Dunham, I. et.al.	Minus	24827522-24827428
	338689	Dunham, I. et.al.	Minus	24893073-24892972
	338695	Dunham, I. et.al.	Minus	25104153-25104016
50	338825	Dunham, I. et.al.	Minus	27664798-27664712
	338842	Dunham, I. et.al.	Minus	27824238-27824079
	338893	Dunham, I. et.al.	Minus	28491807-28491631
	338904	Dunham, I. et.al.	Minus	28766345-28766253
	338935	Dunham, I. et.al.	Minus	29071537-29071461
55	339022	Dunham, I. et.al.	Minus	30523414-30523289
	339034	Dunham, I. et.al.	Minus	30621603-30621422
	339190	Dunham, I. et.al.	Minus	32403103-32402985
	339212	Dunham, I. et.al.	Minus	32494335-32494210
	339213	Dunham, I. et.al.	Minus	32496590-32496440
	339216	Dunham, I. et.al.	Minus	32504250-32504109
60	339233	Dunham, I. et.al.	Minus	32751331-32751238
	339258	Dunham, I. et.al.	Minus	32934756-32934615
	339262	Dunham, I. et.al.	Minus	32971258-32971090
	339263	Dunham, I. et.al.	Minus	32974634-32974452
	339265	Dunham, I. et.al.	Minus	32975943-32975806
65	339338	Dunham, I. et.al.	Minus	33468728-33468606
	339396	Dunham, I. et.al.	Minus	34017306-34017205
	339400	Dunham, I. et.al.	Minus	34045024-34044940
	339425	Dunham, I. et.al.	Minus	34407911-34407798
	325207	6552430	Plus	140049-140170

	329568	3962490	Plus	36331-36750
	329517	3983513	Minus	53197-53269
	325313	5866865	Minus	27385-28192
5	325327	5866875	Plus	75189-75264
	325317	5866878	Minus	156551-156649
	325257	5866895	Plus	10867-10955
	329632	6729060	Plus	192813-193017
	325371	5866920	Minus	1035422-1035536
10	325375	5866920	Minus	1165503-1165810
	325378	5866920	Minus	1187981-1188167
	325469	6017034	Plus	286823-286991
	325470	6017034	Plus	287578-287663
	325576	6552443	Minus	137769-137894
15	325505	6682451	Minus	240852-240946
	325543	6682452	Plus	151873-152057
	329635	5302817	Minus	62522-62622
	329636	5302817	Minus	64969-65078
	325593	5866992	Minus	469726-469860
20	325675	5867014	Plus	955517-955711
	325704	5867028	Plus	156198-156387
	325682	6138923	Plus	370618-370763
	325785	6381957	Plus	61849-62003
	325666	6469822	Plus	16769-16857
25	325818	6682490	Minus	120278-120559
	329777	6002090	Minus	191389-191479
	329768	6015501	Plus	118315-118422
	329759	6048280	Minus	37647-37730
	329731	6065783	Plus	158772-158900
30	329687	6117856	Minus	22165-22288
	329676	6272128	Minus	142207-142359
	329667	6272129	Plus	101355-101745
	329669	6272129	Plus	131223-131291
	329670	6272129	Plus	131351-131495
35	329641	6468233	Minus	105995-106107
	329791	6469354	Minus	131982-132089
	325826	5867048	Minus	46361-46458
	325829	5867052	Plus	232674-233060
	329888	6067149	Minus	37227-37473
40	329893	6525313	Minus	166123-166791
	329899	6563505	Minus	111058-111783
	325988	5867064	Plus	17349-17606
	325855	5867067	Plus	276141-276251
	325999	5867073	Plus	149115-149192
45	326001	5867073	Plus	155223-155348
	325886	5867087	Plus	194694-194915
	325882	5867087	Minus	8178-8347
	325905	5867104	Plus	78779-78876
	325922	5867122	Minus	329063-329134
50	325937	5867132	Minus	152633-152902
	325960	5867147	Minus	162506-162635
	325961	5867147	Minus	165106-165209
	325838	6552452	Plus	171451-171532
	325839	6552452	Plus	181964-182037
55	325840	6552452	Plus	184380-184547
	325844	6552453	Minus	14188-14332
	325870	6682492	Plus	228209-228297
	329984	4646193	Minus	139780-139890
	329976	4878063	Minus	62584-62691
60	329935	6165200	Minus	69059-69127
	329916	6223624	Plus	36396-37195
	330021	6671889	Plus	120938-121032
	330024	6671908	Minus	1005-1270
	330028	6671908	Minus	30015-30144
	326033	5867178	Plus	37261-37333
65	326036	5867178	Minus	120215-120273
	326056	5867184	Minus	181553-181690
	326116	5867193	Plus	45548-45604
	326122	5867194	Plus	144397-144683
	326138	5867203	Minus	179374-179436

	326145	5867204	Minus	52599-52814
	326180	5867211	Minus	182758-183222
	326201	5867216	Minus	166168-166959
5	326207	5867222	Plus	48139-48219
	326226	5867230	Plus	52644-52705
	326233	5867232	Plus	124788-124863
	326238	5867260	Plus	64282-64338
	326241	5867260	Minus	181648-181916
10	326243	5867261	Plus	123838-123978
	326251	5867263	Minus	82716-82822
	326268	5867267	Plus	122114-122765
	326124	5916395	Plus	407102-407560
	326339	6056311	Minus	164637-165251
15	330049	4567182	Minus	314662-315210
	326358	5867293	Plus	9122-9195
	326365	5867297	Minus	96630-96764
	326379	5867327	Plus	32299-32402
	326382	5867327	Minus	50420-50503
20	326390	5867340	Minus	108814-110592
	326424	5867369	Minus	168329-168409
	326453	5867399	Plus	86222-86423
	326472	5867404	Plus	293739-293940
	326492	5867422	Plus	120768-120991
25	326533	5867441	Minus	532153-532280
	330117	6015201	Minus	7340-7680
	330115	6015202	Plus	11403-11677
	330116	6015202	Plus	12109-12418
	330095	6015278	Plus	15343-15814
30	330096	6015278	Plus	49370-49458
	326644	5867559	Plus	42684-42819
	326713	5867595	Plus	121511-121798
	326745	5867611	Plus	127130-127318
	326752	5867615	Minus	1214-1562
35	326753	5867616	Plus	12454-12511
	326598	5867634	Plus	68955-69014
	326667	6552455	Plus	142311-142441
	326855	6552460	Minus	111390-111463
	326812	6682504	Plus	189811-189941
40	327005	5867664	Plus	610847-610907
	327008	5867664	Plus	928737-928811
	326896	5867680	Minus	12032-12122
	326904	5867684	Minus	9280-9606
	326951	6004446	Plus	193812-193998
45	326941	6004446	Plus	62018-62896
	326943	6004446	Minus	89242-89427
	326928	6456782	Minus	291007-291219
	326958	6469836	Minus	42952-43082
	326959	6469836	Minus	43159-43301
50	327039	6531965	Plus	694486-694998
	327127	6682520	Plus	41925-42083
	330158	6580367	Plus	81966-82456
	327204	5867447	Plus	165135-165239
	327208	5867447	Plus	180805-180864
55	327266	5867462	Minus	82400-82615
	327277	5867473	Minus	165616-165715
	327289	5867481	Plus	49296-49536
	327296	5867492	Plus	7627-8166
	327237	5867544	Minus	59702-59813
	327145	5867548	Minus	40482-40551
60	327333	5902477	Minus	141448-141609
	327335	5902477	Minus	142979-143124
	327343	6017017	Minus	12288-12395
	327350	6249563	Minus	41890-41985
65	327358	6552411	Minus	3802-3950
	327360	6552411	Minus	6255-6422
	327409	5867750	Minus	52949-53011
	327424	5867751	Plus	160442-160598
	327430	5867754	Plus	1320-1403
	327470	5867772	Plus	150910-150973

	327460	6004455	Plus	175245-175343
	327498	6017023	Minus	42178-42283
	327509	6117815	Minus	54882-55053
5	327510	6117815	Minus	56824-56944
	327512	6117815	Plus	176256-176325
	327535	6525279	Plus	19105-19175
	330163	6042042	Minus	20321-20385
	330171	6648220	Plus	110889-111575
10	327579	5867824	Minus	37229-38335
	327672	5867843	Minus	69649-69740
	327629	5867872	Plus	49692-49811
	327640	5867890	Plus	9448-9566
	327649	5867899	Plus	205871-205927
15	327612	6525283	Plus	2747-2924
	327718	6525284	Plus	86123-86186
	327801	5867924	Plus	23239-23348
	327762	5867961	Minus	50303-50439
	327763	5867961	Plus	229347-229476
20	327776	5867964	Minus	164308-164486
	327822	5867968	Minus	168886-169633
	327823	5867968	Minus	170359-170433
	327807	5867968	Plus	33745-33811
	327845	6531962	Plus	193402-193549
25	330228	6013527	Minus	3719-3787
	330190	6165182	Plus	36103-36243
	328122	5868031	Plus	158474-158656
	328132	5868038	Minus	126737-126839
	328159	5868065	Minus	52957-53162
30	328168	5868071	Plus	60321-60479
	328175	5868073	Plus	208-271
	328217	5868096	Minus	3742-4362
	327865	5868130	Plus	61503-62205
	327866	5868131	Minus	2893-3046
35	327870	5868131	Plus	53558-53757
	327879	5868142	Minus	77722-77793
	327902	5868158	Minus	133339-133467
	327918	5868165	Plus	547530-547591
	327934	5868184	Plus	41830-42036
40	327959	5868210	Minus	46497-46682
	327976	5868212	Minus	349301-349409
	328020	5902482	Minus	556386-556652
	328042	5902482	Minus	1985085-1986626
	328008	5902482	Plus	296663-297151
45	330301	2905862	Minus	4420-5781
	330299	2905881	Minus	1020-1382
	328274	5868219	Minus	31244-31439
	328595	5868224	Plus	148738-148967
	328591	5868227	Minus	237647-237726
50	328668	5868254	Minus	10888-10984
	328677	5868256	Minus	58708-58950
	328687	5868262	Plus	624479-624585
	328706	5868270	Plus	165501-165614
	328711	5868271	Minus	97797-97990
55	328730	5868289	Plus	8068-8214
	328732	5868289	Plus	37437-37550
	328734	5868289	Plus	50559-50747
	328752	5868298	Minus	114911-115087
	328755	5868301	Minus	145959-146446
60	328761	5868302	Minus	239308-239412
	328775	5868309	Plus	12845-12920
	328784	5868309	Minus	74523-74604
	328787	5868309	Plus	135772-135963
	328809	5868327	Plus	91792-91849
65	328829	5868337	Plus	36309-36630
	328280	5868352	Plus	160563-160631
	328311	5868371	Minus	170560-170826
	328318	5868373	Plus	414945-415620
	328323	5868373	Minus	1080089-1080235
	328348	5868383	Minus	260272-260379

	328377	5868390	Plus	16947-17023
	328436	5868417	Plus	203760-203904
	328504	5868471	Plus	47064-47217
5	328506	5868471	Plus	60716-60830
	328522	5868477	Plus	1972307-1972452
	328525	5868482	Plus	12387-14313
	328541	5868486	Plus	130956-131050
	328662	6004473	Plus	1184773-1184855
10	328663	6004473	Plus	1185279-1186634
	328803	6004475	Minus	291716-291948
	328304	6004478	Minus	3884-3952
	328927	5868500	Minus	428829-428893
	328936	5868500	Minus	1352202-1352259
15	328939	6004481	Minus	131139-131320
	328941	6456765	Minus	9817-9885
	328948	6456765	Plus	28227-28413
	328968	6456775	Plus	117442-118283
	330316	6007576	Minus	119761-119931
20	330350	3056622	Minus	26413-26820
	330351	3056622	Minus	27522-27614
	330348	4544475	Minus	19855-19962
	329034	5868561	Minus	32819-32939
	329046	5868569	Plus	18971-19030
	329053	5868574	Plus	426453-426541
25	329186	5868711	Minus	13108-13225
	329237	5868729	Plus	133238-133339
	329276	5868762	Minus	222629-222709
	329333	5868806	Plus	392666-392746
	329376	5868859	Plus	52356-52694
30	329384	5868869	Minus	116524-116662
	329140	6017060	Plus	290842-290905
	329317	6381976	Plus	614823-615209
	329319	6381976	Plus	721390-721470
	329129	6588026	Plus	144569-144712
35	329373	6682537	Minus	38950-39301
	329412	6682553	Minus	68948-69041
	329424	5868879	Plus	362196-362344
	329446	5868886	Plus	84776-84899
	329449	5868886	Plus	97697-97771

TABLE 14: shows genes, including expression sequence tags, down-regulated in prostate tumor tissue compared to normal prostate tissue as analyzed using Affymetrix/Eos Hu02 GeneChip array. Shown are the ratios of "average" normal prostate to "average" prostate cancer tissues.

10	Pkey:	Unique Eos probeset identifier number			
	ExAccn:	Exemplar Accession number, Genbank accession number			
	UnigeneID:	Unigene number			
	Unigene Title:	Unigene gene title			
	R1:	Background subtracted normal prostate : prostate tumor tissue			
	Pkey	ExAccn	UnigeneID	Unigene Title	R1
15	331328	AA281133	Hs.88808	ESTs	18.53
	320875	D60641	Hs.131921	ESTs	14.55
	300994	AI251936	Hs.146298	ESTs	12.17
	323461	AA418762	Hs.190044	ESTs	10.55
20	301015	AA947682	Hs.217173	ESTs; Weakly similar to Chain A; Cdc42hs-Gdp Complex [H.sapiens]	10.17
	319419	AA543096	Hs.13648	ESTs; Highly similar to mitogen-induced [M.musculus]	9.2
	323486	C05278	Hs.166800	ESTs; Moderately similar to (PYRUVATE DEHYDROGENASE(LIPOAMIDE)) KINASE ISOZYME 4 PRECURSOR [H.sapiens]	8.87
	324882	AW419080	Hs.250645	ESTs	8
25	330569	U57796	Hs.57679	zinc finger protein 192	7.88
	330126			CH.21_p2 gi 6093735	7.8
	316265	AA737400	Hs.142230	ESTs	7.7
	323045	AA148950	Hs.188836	ESTs	7.64
	320668	R58399	Hs.146217	ESTs	7.4
30	330769	AA465192	Hs.16514	ESTs	7.15
	312614	AI766732	Hs.201194	ESTs	7
	314790	AW341754	Hs.189305	ESTs	6.83
	309979	AW452118	Hs.257533	EST	6.74
	314236	AA743396	Hs.189023	ESTs	6.49
35	329192			CH.X_hs gi 5868716	6.1
	324307	AA627642	Hs.4994	transducer of ERBB2; 2 (TOB2)	5.99
	303685	AW500106		EST cluster (not in UniGene) with exon hit	5.82
	314921	AW452382	Hs.257564	ESTs	5.8
	315840	AA679001	Hs.192221	ESTs	5.68
40	332776	AA034364	Hs.256551	ESTs; Weakly similar to !!!! ALU CLASS B WARNING ENTRY !!!! [H.sapiens]	5.43
	313533	AW298141	Hs.157975	ESTs	5.4
	303494	F30712		EST cluster (not in UniGene) with exon hit	5.35
	317490	AI627358	Hs.148367	ESTs	5.31
	332546	D84454	Hs.21899	solute carrier family 35 (UDP-galactose transporter); member 2	5.25
45	334719			CH22_FGENES.421_30	5.25
	300679	AA813958	Hs.207727	ESTs; Moderately similar to KIAA0071 [H.sapiens]	5.22
	311811	AI625304	Hs.190312	ESTs	5.22
	315310	AW511298	Hs.256067	ESTs	5.19
	312871	H86747	Hs.227602	KIAA1116 protein	5.11
50	324715	AI739168		EST cluster (not in UniGene)	4.97
	313870	AW206435	Hs.146057	ESTs	4.97
	321453	N50080	Hs.117827	ESTs	4.78
	316160	AW197887	Hs.253353	ESTs	4.63
	313633	AA766825		EST cluster (not in UniGene)	4.58
55	315850	AW270550	Hs.116957	ESTs	4.53
	303124	AF161350		EST cluster (not in UniGene) with exon hit	4.46
	323346	AL134932	Hs.143607	ESTs	4.4
	301383	AA913591	Hs.126480	ESTs	4.35
	324513	AW501678	Hs.164577	ESTs	4.28
60	303480	AA331906		EST cluster (not in UniGene) with exon hit	4.25
	323591	AA301270		EST cluster (not in UniGene)	4.22
	313603	AW468119		EST cluster (not in UniGene)	4.2
	317863	AI733395	Hs.129124	ESTs	4.1
	312381	R42049	Hs.195473	ESTs	4.08
65	317514	AW451570	Hs.126850	ESTs	4.03
	319750	AA621606	Hs.117956	ESTs	4.03

	322520	T55958	EST cluster (not in UniGene)	4
	314754	AW026761	Hs.134374 ESTs	4
	316088	AI990652	Hs.208973 ESTs	4
	318473	AI939339	Hs.146883 ESTs	3.96
5	307848	AI364186	EST singleton (not in UniGene) with exon hit	3.95
	300730	AW449204	Hs.257125 ESTs	3.94
	303034	W60843	Hs.31570 ESTs	3.93
	324668	AI679131	Hs.201424 ESTs	3.9
10	324674	AA541323	Hs.115831 ESTs	3.88
	300547	N53442	Hs.143443 ESTs	3.83
	316100	AW203986	Hs.213003 ESTs	3.79
	314801	AA481027	Hs.127336 ESTs; Weakly similar to ORF YGR245c [S.cerevisiae]	3.75
	320856	D59945	EST cluster (not in UniGene)	3.74
15	313188	AI039702	Hs.179573 collagen; type I; alpha 2	3.73
	314187	AA804409	Hs.118920 ESTs	3.73
	311826	AA765470	Hs.122826 ESTs	3.7
	302358	D81150	EST cluster (not in UniGene) with exon hit	3.68
	311441	Z38720	Hs.151014 ESTs	3.66
	321914	AA011603	EST cluster (not in UniGene)	3.59
20	332216	H95082	Hs.102332 EST	3.52
	324771	AA631739	EST cluster (not in UniGene)	3.5
	323691	AA317561	EST cluster (not in UniGene)	3.49
	303525	AW516519	Hs.115130 ESTs	3.47
25	309709	AW242630	EST singleton (not in UniGene) with exon hit	3.46
	300038		AFFX control: MurlL4	3.38
	316526	AI088192	Hs.135474 ESTs; Weakly similar to ATP-DEPENDENT RNA HELICASE A [H.sapiens]	3.36
	313029	AA731520	Hs.170504 ESTs	3.35
	304356	AA196027	Hs.195188 glyceraldehyde-3-phosphate dehydrogenase	3.34
30	314610	AI948688	Hs.191805 ESTs	3.33
	329815		CH.14_p2 gjl6624888	3.32
	314949	AI745387	Hs.239124 ESTs	3.31
	300598	N53574	Hs.158932 ESTs	3.3
	329218		CH.X_hs gjl5868726	3.28
35	315706	AW440742	Hs.155556 ESTs	3.28
	303751	AW503637	EST cluster (not in UniGene) with exon hit	3.25
	307783	AI347274	EST singleton (not in UniGene) with exon hit	3.25
	321414	AA324975	Hs.128993 ESTs; Weakly similar to KIAA0465 protein [H.sapiens]	3.25
	312187	AA700439	Hs.188490 ESTs	3.25
40	334061		CH22_FGENES.327_14	3.23
	336036		CH22_FGENES.678_7	3.23
	321477	H67818	Hs.222059 ESTs	3.21
	315760	AW139383	Hs.245437 ESTs	3.2
	316733	AA811713	Hs.163222 ESTs	3.2
45	300855	AW235248	Hs.79828 ESTs	3.2
	323611	AA304986	Hs.145704 ESTs	3.19
	314138	AA740616	EST cluster (not in UniGene)	3.17
	316774	AA814859	EST cluster (not in UniGene)	3.16
	308884	AI833131	Hs.179100 ESTs	3.11
50	331317	AA258222	Hs.87757 ESTs	3.1
	317221	AI989538	Hs.191074 ESTs	3.08
	316386	AA749062	Hs.180285 ESTs	3.08
	321040	H26953	EST cluster (not in UniGene)	3.08
	308828	AI824829	EST singleton (not in UniGene) with exon hit	3.08
55	300778	AA236233	Hs.188716 ESTs	3.07
	316667	AW015940	Hs.232234 ESTs	3.07
	324614	AW503101	EST cluster (not in UniGene)	3.07
	316468	AW293046	Hs.255158 ESTs	3.07
	300671	AI239706	Hs.189886 ESTs	3.06
60	314301	AW297967	Hs.188181 ESTs	3.05
	312335	AW043620	Hs.236993 ESTs	3.03
	322957	AA247755	EST cluster (not in UniGene)	3.01
	316848	AA830053	Hs.126798 ESTs	3.01
	313473	AA009660	Hs.251948 ESTs; Moderately similar to T07D3.7 [C.elegans]	2.99
65	318518	T27119	EST cluster (not in UniGene)	2.98
	313383	AI076370	Hs.134037 ESTs	2.97
	331389	AA458637	Hs.152207 ESTs	2.96
	304257	AA053294	EST singleton (not in UniGene) with exon hit	2.95
	309917	AW340014	EST singleton (not in UniGene) with exon hit	2.95
	319661	H08035	Hs.21398 ESTs; Moderately similar to PUTATIVE GLUCOSAMINE-6-PHOSPHATE	2.95

			ISOMERASE [H.sapiens]	2.95
	321253	AI699484	EST cluster (not in UniGene)	2.93
	321193	AA149508	Hs.103288 ESTs	2.93
	332864		CH22_FGENES.28_4	2.92
5	300027			
		M11507	AFFX control: transferrin receptor	2.91
	324330	AA884766	EST cluster (not in UniGene)	2.88
	320014	AA137114	Hs.170291 ESTs	2.88
10	333916		CH22_FGENES.296_5	2.88
	318885	Z43272	EST cluster (not in UniGene)	2.87
	318146	AI040125	Hs.150521 ESTs	2.87
	323348	AA233056	Hs.191518 ESTs	2.85
	305703	AA825148	Hs.21229 F-box protein Fbw1b	2.84
	335862		CH22_FGENES.629_7	2.83
15	317672	AW205409	Hs.127748 ESTs	2.82
	323416	AI610397	Hs.159560 ESTs	2.81
	312652	AI419909	Hs.160994 ESTs	2.81
	324094	AA382603	EST cluster (not in UniGene)	2.81
	319761	R84237	EST cluster (not in UniGene)	2.8
20	317013	AA864468	Hs.135646 ESTs	2.8
	317383	AA913887	Hs.126511 ESTs	2.78
	314659	AW277121	Hs.254881 ESTs	2.78
	312479	AI950844	Hs.128738 ESTs; Weakly similar to non-lens beta gamma-crystallin like protein [H.sapiens]	2.77
	332808		CH22_FGENES.7_10	2.75
25	311824	AW293826	Hs.250610 ESTs	2.75
	321992	C06003	Hs.116456 ESTs	2.73
	316074	AW517542	Hs.208382 ESTs	2.73
	309839	AW296076	EST singleton (not in UniGene) with exon hit	2.73
	312071	AA683529	Hs.143119 ESTs	2.73
30	312684	AW294020	Hs.117721 ESTs	2.72
	332668	AA062971	Hs.181161 ESTs; Weakly similar to INHIBITOR OF APOPTOSIS PROTEIN 1 [M.musculus]	2.72
	322139	H53744	EST cluster (not in UniGene)	2.72
	304168	H77679	EST singleton (not in UniGene) with exon hit	2.72
	325602		CH.13_hs gil5866994	2.71
35	319885	R59096	Hs.136698 ESTs	2.71
	300611	N75450	EST cluster (not in UniGene) with exon hit	2.71
	316854	AA831215	Hs.159066 ESTs; Weakly similar to predicted using Genefinder [C.elegans]	2.69
	318208	AI091458	Hs.134559 ESTs	2.68
	331623	R38715	Hs.153529 Homo sapiens clone 24540 mRNA sequence	2.68
40	324616	AI823999	Hs.162000 ESTs	2.68
	304968	AA614308	EST singleton (not in UniGene) with exon hit	2.67
	314912	AI431345	Hs.161784 ESTs	2.67
	300767	AW193466	Hs.136525 ESTs	2.67
	313463	AI057369	Hs.122536 ESTs	2.65
45	320600	AA135565	Hs.250739 ESTs	2.65
	301180	AI308989	Hs.156939 ESTs	2.65
	324825	AA704457	Hs.255738 ESTs; Moderately similar to gag [H.sapiens]	2.65
	300336	AW292417	Hs.255074 ESTs; Moderately similar to high-risk human papilloma viruses E6 oncoproteins targeted protein E6TP1 alpha [H.sapiens]	2.64
50	317850	N29974	EST cluster (not in UniGene)	2.64
	339047		CH22_DA59H18.GENSCAN.28-7	2.64
	324580	AA492588	EST cluster (not in UniGene)	2.63
	321142	AI817933	Hs.209584 ESTs	2.62
	319478	R06841	EST cluster (not in UniGene)	2.62
55	300793	AI248571	Hs.186837 ESTs	2.61
	313733	AA836116	EST cluster (not in UniGene)	2.6
	326505		CH.19_hs gil5867435	2.6
	314987	AW015506	Hs.130730 ESTs	2.6
	303114	AF090948	EST cluster (not in UniGene) with exon hit	2.59
60	318709	H24244	Hs.240763 ESTs; Weakly similar to /prediction	2.58
	312878	AI209108	Hs.143946 ESTs	2.57
	329224		CH.X_hs gil5868728	2.56
	328018		CH.06_hs gil5902482	2.56
	323231	AA324437	Hs.177230 ESTs	2.55
65	312887	AW157377	Hs.132910 ESTs	2.55
	315183	AW136134	Hs.220277 ESTs	2.55
	300259	AI479011	Hs.170783 ESTs	2.54
	313240	AI743261	Hs.131860 ESTs	2.54
	316697	AW293174	Hs.252627 ESTs	2.53

	313966	AI807551	Hs.189061	ESTs	2.53
	331263	AA015718		ze31a12.s1 Soares retina N2b4HR Homo sapiens cDNA clone IMAGE:36574 3', mRNA sequence	2.51
5	310683	AW055233	Hs.160870	ESTs	2.5
	302566	AA085996	Hs.248572	Human PAC clone DJ404F18 from Xq23	2.5
	302697	AJ001408		EST cluster (not in UniGene) with exon hit	2.5
	308362	AI613519		EST singleton (not in UniGene) with exon hit	2.49
	322347	AF086538		EST cluster (not in UniGene)	2.49
10	316240	AA974253	Hs.120319	ESTs	2.49
	323208	AA203415	Hs.136200	ESTs	2.48
	321643	W76005	Hs.32094	ESTs	2.48
	330723	AA243617	Hs.31082	ESTs; Highly similar to db83 [R.norvegicus]	2.48
	323455	AA256675	Hs.200438	ESTs; Weakly similar to atypical PKC specific binding protein [R.norvegicus]	2.47
	308383	AI624497		EST singleton (not in UniGene) with exon hit	2.47
15	328744			CH.07_hs gi 5868290	2.47
	332344	W45574	Hs.252497	ESTs	2.47
	328121			CH.06_hs gi 5868031	2.47
	321915	AI670955	Hs.200151	ESTs	2.46
20	314954	AA521381	Hs.187726	ESTs	2.45
	302821	AA188868	Hs.173933	ESTs; Weakly similar to NUCLEAR FACTOR 1/X [H.sapiens]	2.45
	329454			CH.Y_hs gi 5868887	2.45
	336605			CH22_FGENES.420_4	2.45
	300664	AI444628	Hs.256809	ESTs	2.44
25	323362	AL135067	Hs.117182	ESTs	2.44
	300024	M10098		AFFX control: 18S ribosomal RNA	2.44
	325026	AI671168	Hs.12285	ESTs	2.43
	324510	AI148353	Hs.120849	ESTs	2.43
	313389	AI765182	Hs.119903	ESTs	2.43
	301309	M78276	Hs.255917	ESTs	2.43
30	313570	AA041455	Hs.209312	ESTs	2.43
	316504	AW135854	Hs.132458	ESTs	2.42
	319401	R01342		EST cluster (not in UniGene)	2.42
	312827	AI744361	Hs.205591	ESTs; Weakly similar to zinc finger protein Png-1 [M.musculus]	2.42
35	327871			CH.06_hs gi 5868131	2.41
	337173			CH22_FGENES.565-3	2.41
	302948	AA465635		EST cluster (not in UniGene) with exon hit	2.41
	324303	AL118754		EST cluster (not in UniGene)	2.4
	315527	AI791138	Hs.116768	ESTs	2.4
40	315979	AA830515	Hs.222917	ESTs	2.4
	331310	AA253351	Hs.44439	STAT induced STAT inhibitor-4	2.4
	321095	AA017595	Hs.32844	ESTs	2.4
	308561	AI701559		EST singleton (not in UniGene) with exon hit	2.39
	313035	N36417	Hs.144928	ESTs	2.37
	322114	AA643791	Hs.191740	ESTs	2.37
45	313671	W49823	Hs.145553	ESTs	2.37
	303211	AA099548	Hs.191436	ESTs; Highly similar to dJ1118D24.4 [H.sapiens]	2.37
	301256	AA932948		EST cluster (not in UniGene) with exon hit	2.36
	338165			CH22_EM:AC005500.GENSCAN.212-3	2.36
	324692	AA557952		EST cluster (not in UniGene)	2.35
50	318587	AA779704	Hs.168830	ESTs	2.35
	312378	R41582	Hs.109219	retinal degeneration B beta	2.35
	318625	T48446	Hs.193162	ESTs	2.35
	305181	AA663726	Hs.116922	EST	2.35
	300815	AA286678		EST cluster (not in UniGene) with exon hit	2.34
55	324063	AW292740	Hs.254815	ESTs	2.34
	315859	AA682305	Hs.133268	ESTs	2.33
	305092	AA642912		EST singleton (not in UniGene) with exon hit	2.33
	306598	AI000320		EST singleton (not in UniGene) with exon hit	2.33
	300307	AI651016	Hs.246311	ESTs	2.33
60	321348	Z49979		EST cluster (not in UniGene)	2.33
	325112	AI903770	Hs.124344	ESTs	2.32
	336679			CH22_FGENES.43-7	2.32
	321383	AJ002574		EST cluster (not in UniGene)	2.32
	337357			CH22_FGENES.730-6	2.31
65	300680	AW468066	Hs.257712	ESTs; Weakly similar to KIAA0986 protein [H.sapiens]	2.31
	327120			CH.21_hs gi 6531970	2.31
	302761	AW250553		EST cluster (not in UniGene) with exon hit	2.3
	312132	AI475490	Hs.170577	ESTs	2.3
	315639	AA827652		EST cluster (not in UniGene)	2.3

	312189	T95594	Hs.187435	ESTs	2.3
	306537	AA991705		EST singleton (not in UniGene) with exon hit	2.3
	327061			CH.21_hs gi 6531965	2.3
5	315391	AA759098	Hs.192007	ESTs	2.3
	322384	AI968646	Hs.33862	ESTs	2.29
	323206	AA203339	Hs.220750	ESTs	2.29
	318110	AI680915	Hs.201379	ESTs	2.28
	335250			CH22_FGENES.516_11	2.28
10	331696	Z38907	Hs.91662	KIAA0888 protein	2.28
	318327	AW294013	Hs.200942	ESTs	2.28
	324980	AA969121	Hs.254296	ESTs	2.28
	319429	AI608881	Hs.11482	ESTs; Highly similar to junctional adhesion molecule [H.sapiens]	2.28
	310601	AI970543	Hs.192605	ESTs	2.28
15	318905	Z43395		EST cluster (not in UniGene)	2.28
	323442	AA252753	Hs.164039	ESTs	2.27
	304428	AA342250	Hs.99819	ubiquitin specific protease 16	2.27
	313352	AW292127	Hs.144758	ESTs	2.27
	316491	AA766025	Hs.238794	EST	2.27
20	317751	AI697668	Hs.202241	ESTs	2.26
	314136	AA229781	Hs.221962	ESTs	2.26
	306665	AI004614	Hs.130577	EST	2.26
	303946	AW474196	Hs.221604	ESTs	2.25
	313435	AA769123		EST cluster (not in UniGene)	2.25
25	317679	AA968799	Hs.150289	ESTs	2.25
	322370	AA330095		EST cluster (not in UniGene)	2.25
	306620	AI000929		EST singleton (not in UniGene) with exon hit	2.24
	329109			CH.X_hs gi 5868626	2.24
	311043	AI871209	Hs.177128	ESTs	2.24
30	300228	AI458372	Hs.158748	ESTs; Weakly similar to synapsin Ib [M.musculus]	2.24
	307223	AI193698	Hs.184776	ribosomal protein L23a	2.24
	309023	AI888045		EST singleton (not in UniGene) with exon hit	2.23
	310749	AI493675	Hs.170332	ESTs	2.23
	316769	AI914939	Hs.212184	ESTs	2.22
35	320409	AA356195		EST cluster (not in UniGene)	2.21
	333149			CH22_FGENES.87_8	2.21
	324951	M86125	Hs.137487	ESTs	2.21
	321939	AI791617	Hs.145068	ESTs	2.2
	320594	AI863952	Hs.169436	arginyltransferase 1	2.2
40	320722	R67430	Hs.172787	ESTs	2.2
	321781	D78667		EST cluster (not in UniGene)	2.2
	328903			CH.08_hs gi 5868514	2.2
	303889	T19204		EST cluster (not in UniGene) with exon hit	2.2
	325045	T08845		EST cluster (not in UniGene)	2.2
45	312828	AI865455	Hs.211818	ESTs; Moderately similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]	2.19
	335109			CH22_FGENES.494_15	2.18
	330878	AA131471	Hs.71440	ESTs	2.18
	311289	AI971362	Hs.231945	ESTs	2.18
	304608	AA513456		EST singleton (not in UniGene) with exon hit	2.18
50	337393			CH22_FGENES.747-4	2.18
	332812			CH22_FGENES.7_14	2.18
	327665			CH.04_hs gi 5867839	2.18
	314581	AW504859	Hs.237849	ESTs	2.17
	326508			CH.19_hs gi 6682496	2.17
55	301242	AW161535	Hs.258803	ESTs	2.17
	312780	AI765651	Hs.172900	ESTs	2.17
	315954	AW276810	Hs.254859	ESTs	2.16
	311179	AI880843	Hs.223333	ESTs	2.16
	315320	AI084182	Hs.186895	ESTs	2.16
	313017	AI015203	Hs.118015	ESTs	2.16
60	312430	AW139117	Hs.117494	ESTs	2.15
	300864	AA406539	Hs.190958	ESTs	2.15
	314753	AA463262		EST cluster (not in UniGene)	2.15
	322574	AF156548		EST cluster (not in UniGene)	2.15
65	321409	C03864		EST cluster (not in UniGene)	2.15
	321205	AA002047		EST cluster (not in UniGene)	2.14
	320406	AA353895	Hs.152983	HUS1 (S. pombe) checkpoint homolog	2.14
	337646			CH22_EM:AC000097.GENSCAN.11-2	2.13
	303084	AF174008		EST cluster (not in UniGene) with exon hit	2.13
	312185	AA654772	Hs.186564	ESTs	2.13

	306813	AI066544	EST singleton (not in UniGene) with exon hit	2.13	
	314465	AA602917	Hs.156974 ESTs	2.12	
	318168	AI821782	Hs.220587 ESTs; Moderately similar to !!!! ALU SUBFAMILY SC WARNING ENTRY !!!! [H.sapiens]	2.12	
	315990	AI800041	Hs.190555 ESTs	2.11	
5	320712	R66867	EST cluster (not in UniGene)	2.11	
	318487	AI167877	Hs.143716 ESTs	2.11	
	317462	AW015206	Hs.178784 ESTs	2.11	
	304384	AA235482	Hs.62954 ferritin; heavy polypeptide 1	2.11	
	314544	AA399018	Hs.250835 ESTs	2.1	
10	319881	T72744	EST cluster (not in UniGene)	2.1	
	328078		CH.06_hs gjl5868008	2.1	
	317354	AW090770	Hs.192271 ESTs	2.1	
	308617	AI738720	EST singleton (not in UniGene) with exon hit	2.09	
	311568	AW439969	Hs.218177 ESTs	2.09	
15	313605	AI761786	Hs.204674 ESTs	2.09	
	314289	AA848118	Hs.221216 ESTs	2.08	
	332933		CH22_FGENES.38_7	2.08	
	325498		CH.12_hs gjl5866967	2.08	
	313659	AW296067	Hs.124106 ESTs	2.08	
20	324596	AW149321	Hs.105411 ESTs	2.08	
	324783	AA640770	EST cluster (not in UniGene)	2.07	
	302696	AA347452	EST cluster (not in UniGene) with exon hit	2.07	
	313418	AW450674	Hs.114696 ESTs	2.06	
	326920		CH.21_hs gjl6456782	2.06	
25	327574		CH.03_hs gjl5867818	2.06	
	323207	AI052795	Hs.192201 ESTs	2.06	
	303753	AW503733	Hs.170315 ESTs	2.05	
	305235	AA670480	EST singleton (not in UniGene) with exon hit	2.05	
	316055	AA693880	EST cluster (not in UniGene)	2.05	
30	317194	AW445167	Hs.126036 ESTs	2.05	
	319565	AW408683	Hs.32922 ESTs	2.05	
	335146		CH22_FGENES.499_2	2.05	
	301475	AI678183	Hs.170917 prostaglandin E receptor 3 (subtype EP3)	2.04	
	312442	AA120970	Hs.143199 ESTs	2.04	
35	322502	R62925	Hs.243665 ESTs	2.04	
	303693	AA290875	Hs.30120 ESTs	2.04	
	310179	AI215643	Hs.171381 ESTs	2.03	
	321121	W23285	EST cluster (not in UniGene)	2.03	
	331330	AA282197	Hs.89002 ESTs; Highly similar to CGI-07 protein [H.sapiens]	2.03	
40	306557	AA994530	EST singleton (not in UniGene) with exon hit	2.03	
	317865	AI298794	Hs.129130 ESTs	2.03	
	318667	AI493742	Hs.165210 ESTs	2.02	
	318042	AW294522	Hs.149991 ESTs	2.02	
	323818	AW245528	Hs.134754 ESTs	2.02	
45	331286	AA137062	Hs.103853 ESTs	2.01	
	311262	AI989942	Hs.232150 ESTs	2.01	
	335601		CH22_FGENES.581_41	2.01	
	311351	AI682303	Hs.201274 ESTs	2.01	
	312996	AA249018	EST cluster (not in UniGene)	2.01	
50	328190		CH.06_hs gjl5868077	2	
	338030		CH22_EM:AC005500.GENSCAN.148-16	2	
	333940		CH22_FGENES.301_6	2	
	328227		CH.06_hs gjl5868105	2	
	331481	N27448	Hs.43944 EST	2	
55	335288		CH22_FGENES.527_1	2	
	307513	AI274307	EST singleton (not in UniGene) with exon hit	2	
	323316	AL134620	EST cluster (not in UniGene)	2	
	319479	R21945	Hs.256153 ESTs	2	
	303482	AA502583	Hs.197271 ESTs	2	
60	327489		CH.02_hs gjl6004459	1.99	
	323935	AW175841	Hs.192183 ESTs	1.99	
	309575	AW168096	Hs.195188 glyceraldehyde-3-phosphate dehydrogenase	1.99	
	337043		CH22_FGENES.439-19	1.98	
	312897	AI828174	Hs.227049 ESTs	1.98	
65	307881	AI370434	EST singleton (not in UniGene) with exon hit	1.98	
	328656		CH.07_hs gjl6004473	1.98	
	314569	AA813784	Hs.123001 ESTs	1.98	
	332783	W45302	Hs.87889 helicase-mol	1.98	
	315259	AA701499	Hs.148115 ESTs	1.98	

	313171	N67879	Hs.157695	ESTs	1.97
	318060	AI241421	Hs.132236	ESTs	1.97
	332256	N66393	Hs.102754	ESTs	1.97
	312110	AI962180	Hs.226803	ESTs	1.97
5	335864		CH22_FGENES.629_9		1.97
	320389	W00545	Hs.171785	ESTs	1.97
	314065	AA868267	Hs.85524	ESTs	1.96
	323086	H15474	Hs.12214	Homo sapiens clone 23716 mRNA sequence	1.96
	323919	AA862973	Hs.220704	ESTs	1.96
10	310750	AI373163	Hs.170333	ESTs	1.96
	309435	AW090537		EST singleton (not in UniGene) with exon hit	1.96
	300129	AW028820		EST cluster (not in UniGene) with exon hit	1.96
	320130	AI820675	Hs.203804	ESTs	1.95
	323787	AW373446	Hs.169885	ESTs; Weakly similar to cDNA EST EMBL:T02216 comes from this gene [C.elegans]	1.95
15	338112		CH22_EM:AC005500.GENSCAN.185-24		1.95
	313625	AW468402	Hs.254020	ESTs	1.95
	325240		CH.10_hs gi 5866848		1.95
	331833	AA412102	Hs.250911	interleukin 13 receptor; alpha 1	1.95
	332252	N63882		za21f9.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:293225 3', mRNA sequence	1.95
20	300279	AW237425	Hs.253817	ESTs	1.95
	326023		CH.17_hs gi 5867245		1.95
	321609	H86021	Hs.198800	ESTs; Weakly similar to hMmTRA1b [H.sapiens]	1.94
	324183	AA402453	Hs.113011	ESTs	1.94
25	336276		CH22_FGENES.762_5		1.94
	334913		CH22_FGENES.456_3		1.94
	325417		CH.12_hs gi 5866925		1.94
	318489	AW043590	Hs.225023	ESTs	1.94
	318455	AI148763		EST cluster (not in UniGene)	1.94
30	306890	AI092235		EST singleton (not in UniGene) with exon hit	1.94
	315073	AW452948	Hs.257631	ESTs	1.94
	321289	R84687	Hs.226306	ESTs	1.94
	308521	AI689808		EST singleton (not in UniGene) with exon hit	1.93
	306382	AA968967		EST singleton (not in UniGene) with exon hit	1.93
35	331320	AA262999	Hs.42788	ESTs	1.93
	324279	AA501412	Hs.191688	ESTs; Weakly similar to Pro-Pol-dUTPase polyprotein [M.musculus]	1.93
	309577	AW168753		EST singleton (not in UniGene) with exon hit	1.93
	327014		CH.21_hs gi 5867664		1.93
	303488	AW025860		EST cluster (not in UniGene) with exon hit	1.93
40	306561	AA995223	Hs.129559	EST	1.92
	330694	AA019806	Hs.108447	spinocerebellar ataxia 7 (olivopontocerebellar atrophy with retinal degeneration)	1.92
	313083	N50545	Hs.159200	ESTs	1.92
	327752		CH.05_hs gi 5867949		1.92
	318674	AA295490		EST cluster (not in UniGene)	1.92
45	301267	AW297762	Hs.255690	ESTs	1.91
	332092	AA608787	Hs.112590	ESTs	1.91
	323509	AL036947		EST cluster (not in UniGene)	1.91
	321452	AA317554		EST cluster (not in UniGene)	1.91
	311483	AI765013	Hs.209128	ESTs	1.91
50	300976	AI246374	Hs.185861	ESTs	1.91
	323715	AA322155		EST cluster (not in UniGene)	1.91
	313800	AW296132	Hs.166674	ESTs	1.91
	332029	AA489697	Hs.145053	ESTs	1.91
	304013	AW518573	Hs.156110	Immunoglobulin kappa variable 1D-8	1.91
55	322019	AA354549	Hs.41181	Homo sapiens mRNA; cDNA DKFZp727C191 (from clone DKFZp727C191)	1.91
	334150		CH22_FGENES.339_1		1.9
	310094	AW450967	Hs.235240	ESTs	1.9
	316218	AW207642	Hs.174021	ESTs	1.9
	324774	AI031771	Hs.132586	ESTs	1.9
60	326507		CH.19_hs gi 5867435		1.9
	314570	AA405696		EST cluster (not in UniGene)	1.9
	336268		CH22_FGENES.758_2		1.9
	315278	AI985544	Hs.116429	ESTs	1.9
	325824		CH.15_hs gi 5867048		1.9
65	316277	AA737780	Hs.213392	ESTs	1.9
	323181	AA418583	Hs.143621	ESTs	1.9
	301438	AA961643	Hs.127716	ESTs	1.89
	307050	AI147341	Hs.146734	EST	1.89
	306830	AI075803		EST singleton (not in UniGene) with exon hit	1.89

	302426	AL049925	Hs.225984	DKFZP547G0910 protein	1.89
	320127	H72615	Hs.17268	ESTs	1.89
	337736			CH22_EM:AC000097.GENSCAN.100-2	1.89
5	331319	AA262755	Hs.194264	ESTs	1.88
	310767	AI377505	Hs.158835	ESTs	1.88
	314880	AI732169	Hs.105429	ESTs	1.88
	312539	AI004377	Hs.200360	ESTs	1.88
	309674	AW205604	Hs.168034	ESTs; Weakly similar to !!!! ALU SUBFAMILY SP WARNING ENTRY !!!! [H.sapiens]	1.88
	314621	AI627478	Hs.187670	ESTs	1.88
10	319495	AI972146	Hs.192756	ESTs	1.88
	313472	AA007374		EST cluster (not in UniGene)	1.88
	302705	U09060		EST cluster (not in UniGene) with exon hit	1.88
	329511			CH.10_p2 gi 3983514	1.88
	317140	AI699412	Hs.201925	ESTs	1.87
15	302598	AI815985	Hs.129683	ubiquitin-conjugating enzyme E2D 1 (homologous to yeast UBC4/5)	1.87
	301153	AA725670	Hs.120485	ESTs; Weakly similar to serine/threonine kinase with SH3 domain; leucine zipper domain and proline rich domain [H.sapiens]	1.87
	332222	N28271	Hs.176618	ESTs	1.87
20	330703	AA055475	Hs.104143	clathrin; light polypeptide (Lca)	1.87
	318470	AI159863	Hs.143713	ESTs	1.87
	314014	AW291847	Hs.121715	ESTs; Weakly similar to HP protein [H.sapiens]	1.87
	300370	AI827817		EST cluster (not in UniGene) with exon hit	1.86
	312329	R84768	Hs.13399	Homo sapiens clone 25032 mRNA sequence	1.86
	325587			CH.12_hs gi 6682462	1.86
25	310237	AI884313	Hs.158906	ESTs	1.86
	318872	R13085		EST cluster (not in UniGene)	1.86
	303431	AA317915		EST cluster (not in UniGene) with exon hit	1.86
	338427			CH22_EM:AC005500.GENSCAN.349-1	1.86
	300452	AI352293	Hs.191098	ESTs	1.85
30	321279	H85330	Hs.146060	ESTs	1.85
	301690	F05865	Hs.249180	ubiquitin-conjugating enzyme E2E 2 (homologous to yeast UBC4/5)	1.85
	307932	AJ230822		EST singleton (not in UniGene) with exon hit	1.85
	318292	AI679966	Hs.150603	ESTs	1.85
	310254	AI239811	Hs.157491	ESTs	1.85
35	311790	AW016437	Hs.233462	ESTs	1.84
	314248	AA278347	Hs.126078	ESTs	1.84
	335586			CH22_FGENES.581_25	1.84
	339209			CH22_FF113D11.GENSCAN.6-4	1.84
	307954	AI419692		EST singleton (not in UniGene) with exon hit	1.84
40	302549	AF055136	Hs.248162	tectorin alpha	1.84
	321629	H87213	Hs.158092	ESTs	1.84
	301239	AA807558		EST cluster (not in UniGene) with exon hit	1.84
	332434	N75542	Hs.75356	transcription factor 4	1.84
	327192			CH.01_hs gi 5867445	1.83
45	310214	AI220072	Hs.165893	ESTs	1.83
	320516	R33857	Hs.181479	ESTs; Weakly similar to E-SELECTIN PRECURSOR [H.sapiens]	1.83
	324231	W60827		EST cluster (not in UniGene)	1.83
	336616			CH22_FGENES.613_5	1.83
	328799			CH.07_hs gi 5868316	1.83
50	324661	AW504161		EST cluster (not in UniGene)	1.83
	313190	AA766707	Hs.153039	ESTs	1.83
	301979	L28168	Hs.121495	potassium voltage-gated channel; Isk-related family; member 1	1.82
	302099	AL021397	Hs.137576	ribosomal protein L34 pseudogene 1	1.82
	320187	T99949		EST cluster (not in UniGene)	1.82
55	320791	R78808	Hs.93961	ESTs; Weakly similar to !!!! ALU CLASS A WARNING ENTRY !!!! [H.sapiens]	1.82
	305733	AA829535	Hs.84298	CD74 antigen (invariant polypept of MHC; class II antigen-associated)	1.82
	308280	AI569349	Hs.180920	ribosomal protein S9	1.81
	321533	W78877	Hs.40111	ESTs	1.81
	312946	AI915122	Hs.204087	ESTs; Weakly similar to F33D11.9b [C.elegans]	1.81
60	319474	H90265	Hs.100636	ESTs	1.81
	329519			CH.10_p2 gi 3983510	1.81
	324685	AA220982		EST cluster (not in UniGene)	1.81
	320697	N62937	Hs.139181	ESTs	1.81
	329246			CH.X_hs gi 5868732	1.81
65	332000	AA481271	Hs.193945	ESTs	1.81
	310811	AI420990	Hs.161303	ESTs	1.81
	325866			CH.16_hs gi 5867076	1.81
	322064	Z78343		EST cluster (not in UniGene)	1.8
	333712			CH22_FGENES.251_1	1.8

	313457	AA576052	Hs.193223	ESTs	1.8	
	321591	H85687	Hs.117927	ESTs	1.8	
	330260		CH.05_p2	gil6671884	1.8	
5	311080	AI656320	Hs.197711	ESTs	1.8	
	329522		CH.10_p2	gil3983507	1.8	
	322889	AA081924	Hs.211417	ESTs	1.8	
	300175	AI275011	Hs.204877	ESTs	1.8	
	330976	H20560	Hs.244624	ESTs	1.8	
10	300208	AI341180	Hs.196115	ESTs; Weakly similar to FIBRILLIN 1 PRECURSOR [H.sapiens]	1.79	
	319635	R17531		EST cluster (not in UniGene)	1.79	
	313454	AA730673	Hs.188634	ESTs	1.79	
	303093	AI400310	Hs.148958	ESTs	1.79	
	309815	AW292760		EST singleton (not in UniGene) with exon hit	1.79	
	326506			CH.19_hs	gil5867435	1.79
15	319845	AA649011	Hs.187902	ESTs	1.79	
	300290	AI623739	Hs.186387	ESTs	1.79	
	312180	AI248285	Hs.118348	ESTs	1.79	
	313058	D81015	Hs.125382	ESTs	1.79	
20	330120			CH.19_p2	gil6671864	1.78
	328412			CH.07_hs	gil5868405	1.78
	302345	NM_000565		EST cluster (not in UniGene) with exon hit	1.78	
	308100	AI475949		EST singleton (not in UniGene) with exon hit	1.78	
	311386	AW205705	Hs.207514	ESTs	1.78	
	330282			CH.05_p2	gil6671910	1.78
25	318856	Z43011	Hs.21169	ESTs	1.78	
	312486	AA845630	Hs.117904	ESTs	1.78	
	325450			CH.12_hs	gil5866941	1.78
	321206	H54178	Hs.226469	ESTs	1.78	
30	330977	H20826	Hs.31783	ESTs	1.78	
	303487	AA333666		EST cluster (not in UniGene) with exon hit	1.77	
	310398	AI264671	Hs.164166	ESTs	1.77	
	313230	AI540166	Hs.129563	ESTs	1.77	
	317747	AI683782	Hs.128245	ESTs	1.77	
35	303381	AL038841	Hs.163313	ESTs; Weakly similar to !!!! ALU SUBFAMILY SB WARNING ENTRY !!!! [H.sapiens]	1.77	
	336123			CH22_FGENES.701_8	1.77	
	300185	AI286182	Hs.208484	ESTs	1.77	
	316002	AW451733	Hs.119824	ESTs	1.77	
	319850	AA001811	Hs.83722	ESTs	1.77	
40	329941			CH.16_p2	gil6165199	1.77
	328329			CH.07_hs	gil5868375	1.77
	322934	AI493054	Hs.158968	ESTs	1.77	
	325902			CH.16_hs	gil5867101	1.76
	322239	W01813	Hs.12109	WD40 protein C1a01	1.76	
45	303530	AI274851	Hs.258744	ESTs	1.76	
	300980	AI025527	Hs.222097	ESTs	1.76	
	331909	AA437300	Hs.178210	ESTs	1.76	
	321553	H92449	Hs.116406	ESTs	1.76	
	301618	T52760		EST cluster (not in UniGene) with exon hit	1.76	
50	319592	AA627356	Hs.163315	ESTs	1.76	
	318511	T26528	Hs.227175	ESTs; Weakly similar to !!!! ALU SUBFAMILY SQ WARNING ENTRY !!!! [H.sapiens]	1.76	
	327183			CH.01_hs	gil5867442	1.76
	313516	AA029058	Hs.135145	ESTs	1.76	
	318644	AI752482		EST cluster (not in UniGene)	1.76	
55	321632	AA419617		EST cluster (not in UniGene)	1.76	
	324657	AW451142	Hs.255628	ESTs	1.76	
	300437	AW449374	Hs.257149	ESTs	1.75	
	319775	AA504429	Hs.6211	methyl-CpG binding domain protein 1	1.75	
	314775	AI149880	Hs.188809	ESTs	1.75	
60	337460			CH22_FGENES.780-5	1.75	
	309849	AW297444		EST singleton (not in UniGene) with exon hit	1.75	
	301471	AA995014	Hs.129544	ESTs; Weakly similar to ORF YLL027w [S.cerevisiae]	1.75	
	312739	AI318426	Hs.155925	ESTs	1.75	
	319995	H15355	Hs.60887	ESTs	1.75	
65	326495			CH.19_hs	gil5867423	1.75
	337497			CH22_FGENES.801-4	1.75	
	322633	AA004534	Hs.153981	ESTs	1.75	
	332177	F10812	Hs.101433	ESTs	1.75	
	326930			CH.21_hs	gil6456782	1.75
	316893	AA837332		EST cluster (not in UniGene)	1.75	

	324826	AA704806	Hs.143842	ESTs	1.75
	311289	AI656924	Hs.174257	ESTs	1.75
	309375	AW075342		EST singleton (not in UniGene) with exon hit	1.75
5	314171	AI821895	Hs.193481	ESTs	1.75
	311684	AI990741	Hs.252809	ESTs	1.75
	334387			CH22_FGENES.380_1	1.75
	312195	AI300101	Hs.252222	ESTs	1.75
	315707	AI418055	Hs.161160	ESTs	1.74
	324349	AW501470		EST cluster (not in UniGene)	1.74
10	300724	AI762929	Hs.206134	ESTs; Weakly similar to similar to reverse transcriptase [C.elegans]	1.74
	309906	AW339340		EST singleton (not in UniGene) with exon hit	1.74
	303714	AW501336		EST cluster (not in UniGene) with exon hit	1.74
	318704	Z24981		EST cluster (not in UniGene)	1.74
	303027	AF111178		EST cluster (not in UniGene) with exon hit	1.74
15	322601	W92924		EST cluster (not in UniGene)	1.74
	319382	H93199	Hs.33665	ESTs	1.74
	315858	AA737345		EST cluster (not in UniGene)	1.74
	332243	N55484	Hs.220540	ESTs; Highly similar to ARL HYDROCARBON RECEPTOR NUCLEAR TRANSLOCATOR [H.sapiens]	1.74
20	330951	H02566	Hs.191268	Homo sapiens mRNA; cDNA DKFZp434N174 (from clone DKFZp434N174)	1.74
	324044	AL045752	Hs.211519	ESTs	1.73
	320630	AA199847		EST cluster (not in UniGene)	1.73
	327288			CH.01_hs gij5867481	1.73
	314986	AI201367	Hs.142860	ESTs	1.73
25	319078	H17255	Hs.144515	ESTs	1.73
	326278			CH.17_hs gij5867269	1.73
	302552	H49792		EST cluster (not in UniGene) with exon hit	1.73
	322322	AF086431		EST cluster (not in UniGene)	1.73
	327075			CH.21_hs gij6531965	1.73
30	317392	AI797588	Hs.145459	ESTs	1.73
	300810	AI076890	Hs.186949	ESTs	1.73
	315978	AA830893	Hs.119769	ESTs	1.73
	323903	AA773580	Hs.193598	ESTs	1.73
	330803	AA004699	Hs.150580	putative translation initiation factor	1.73
35	309845	AW296802	Hs.255580	EST	1.73
	314963	AI689617	Hs.200934	ESTs	1.73
	311710	F09774	Hs.175971	ESTs	1.73
	315315	AI984592	Hs.15088	ESTs	1.73
	300378	AA663560	Hs.235873	ESTs; Weakly similar to K11C4.2 [C.elegans]	1.73
40	316141	AW303457		EST cluster (not in UniGene)	1.72
	319826	T71739	Hs.75442	albumin	1.72
	312961	AI033922	Hs.122517	ESTs	1.72
	334379			CH22_FGENES.379_11	1.72
	305854	AA862733		EST singleton (not in UniGene) with exon hit	1.72
45	313031	N34927	Hs.186566	ESTs	1.72
	329728			CH.14_p2 gij6065785	1.72
	312090	N57692	Hs.118064	ESTs	1.72
	323341	AL134875	Hs.192386	ESTs	1.72
	302077	AA310580	Hs.132898	Homo sapiens chromosome 11; BAC CIT-HSP-311e8 (BC269730) containing the hFEN1 gene	1.71
50	310766	AI971438	Hs.158824	ESTs	1.71
	311450	AI809985	Hs.203340	ESTs	1.71
	311792	AW238064	Hs.253909	ESTs	1.71
	321500	H71999		EST cluster (not in UniGene)	1.71
55	311948	T78791	Hs.241569	ESTs; Moderately smlr to !!!! ALU SUBFAMILY SQ WARNING ENTRY !!!! [H.sapiens]	1.71
	302270	R56151		EST cluster (not in UniGene) with exon hit	1.71
	329089			CH.X_hs gij5868614	1.71
	322331	AF086467		EST cluster (not in UniGene)	1.71
	318235	AI080361	Hs.134217	ESTs	1.71
60	304561	AA489792		EST singleton (not in UniGene) with exon hit	1.71
	312681	AI028149	Hs.193124	pyruvate dehydrogenase kinase; isoenzyme 3	1.71
	310250	AI478629	Hs.158465	ESTs	1.71
	338178			CH22_EM:AC005500.GENSCAN.219-6	1.71
	338910			CH22_DJ32110.GENSCAN.11-2	1.71
65	321225	AL080073	Hs.251414	Homo sapiens mRNA; cDNA DKFZp564B1462 (from clone DKFZp564B1462)	1.7
	322289	AA534550	Hs.539	ribosomal protein S29	1.7
	319802	AI701489	Hs.202501	ESTs	1.7
	314022	AW452420	Hs.248678	ESTs	1.7
	314937	AA515602	Hs.152330	ESTs	1.7

	300580	AA761322	Hs.220538	ESTs	1.7
	304398	AA262785		EST singleton (not in UniGene) with exon hit	1.7
	313421	AW339515	Hs.163700	ESTs	1.7
	309763	AW270182		EST singleton (not in UniGene) with exon hit	1.7
5	322092	AF085833		EST cluster (not in UniGene)	1.7
	315603	AA764768	Hs.121158	ESTs	1.7
	325031	T08597		EST cluster (not in UniGene)	1.7
	327157			CH.01_hs gij5866841	1.7
	314809	AI741461	Hs.161904	ESTs	1.7
10	320361	H67220	Hs.146406	nitrilase 1	1.69
	324721	AW402302	Hs.43616	ESTs	1.69
	328624			CH.07_hs gij5868246	1.69
	303344	AA255977	Hs.250646	ESTs; Highly similar to ubiquitin-conjugating enzyme [M.musculus]	1.69
	328960			CH.08_hs gij6456775	1.69
15	315702	AA657501	Hs.146315	ESTs	1.69
	302385	AJ224172	Hs.204096	lipophilin B (uteroglobin family member); prostatein-like	1.68
	319699	R14537		EST cluster (not in UniGene)	1.68
	309506	AW137700		EST singleton (not in UniGene) with exon hit	1.68
	330417	D84424	Hs.57697	hyaluronan synthase 1	1.68
20	315296	AA876905	Hs.125286	ESTs	1.68
	328538			CH.07_hs gij5868485	1.68
	323923	AA354146		EST cluster (not in UniGene)	1.68
	320303	AL079289	Hs.137154	Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 35971	1.68
	302967	AI927068	Hs.110853	ESTs; Weakly similar to R10D12.12 [C.elegans]	1.68
25	310695	AI472124	Hs.157757	ESTs	1.68
	307512	AI273815	Hs.242463	keratin 8	1.68
	338506			CH22_EM:AC005500.GENSCAN.390-10	1.68
	331722	AA195405	Hs.110347	Homo sapiens mRNA for alpha integrin binding protein 80; partial	1.68
	301431	R05385		EST cluster (not in UniGene) with exon hit	1.68
30	318853	Z42977	Hs.21062	ESTs	1.68
	323032	AW244073	Hs.145946	ESTs	1.68
	317538	AW137772	Hs.185980	ESTs	1.68
	325780			CH.14_hs gij6381953	1.67
	321739	AL080280		EST cluster (not in UniGene)	1.67
35	319808	T58960		EST cluster (not in UniGene)	1.67
	313443	AA249037		EST cluster (not in UniGene)	1.67
	331366	AA424754	Hs.43149	ESTs	1.67
	316443	AI797592	Hs.207407	ESTs	1.67
	322878	AA081820		EST cluster (not in UniGene)	1.67
40	330320			CH.08_p2 gij5932415	1.67
	329081			CH.X_hs gij5868602	1.67
	334026			CH22_FGENES.318_3	1.67
	317791	AI801500	Hs.128457	ESTs	1.67
	322235	AF086106		EST cluster (not in UniGene)	1.66
45	331148	R73816	Hs.17385	ESTs	1.66
	325452			CH.12_hs gij5866941	1.66
	315106	AW452184	Hs.232100	ESTs	1.66
	326014			CH.16_hs gij5867160	1.66
	307130	AI185234		EST singleton (not in UniGene) with exon hit	1.66
50	300943	AA524545	Hs.224630	ESTs	1.66
	319402	W21298		EST cluster (not in UniGene)	1.66
	310889	AI457946	Hs.170437	ESTs; Weakly similar to hyperpolarization-activated; cyclic nucleotide-gated channel 2 [H.sapiens]	1.66
	323371	AL135118		EST cluster (not in UniGene)	1.66
55	335568			CH22_FGENES.581_4	1.66
	320654	AW263086	Hs.118112	ESTs	1.66
	338983			CH22_DA59H18.GENSCAN.3-1	1.65
	330002			CH.16_p2 gij6623963	1.65
	315343	AW205477	Hs.179891	ESTs	1.65
60	334487			CH22_FGENES.395_9	1.65
	312169	AI064824	Hs.193385	ESTs	1.65
	309668	AW204480	Hs.253414	EST	1.65
	309518	AW148928	Hs.248895	EST	1.65
	307965	AI421641		EST singleton (not in UniGene) with exon hit	1.65
65	316787	AW369770	Hs.130351	ESTs	1.65
	300835	AA401858	Hs.224843	ESTs	1.65
	338763			CH22_EM:AC005500.GENSCAN.517-16	1.65
	303327	AA232729	Hs.154302	ESTs	1.65
	313231	AW139993	Hs.163682	ESTs	1.65

	334073		CH22_FGENES.327_28	1.65
	319901 T77136	Hs.8765	RNA helicase-related protein	1.65
	326530		CH.19_hs gjl5867441	1.65
5	301126 Al802877	Hs.210843	ESTs; Weakly similar to dJ1039K5.2 [H.sapiens]	1.65
	314043 AA827082		EST cluster (not in UniGene)	1.65
	304387 AA236027		EST singleton (not in UniGene) with exon hit	1.65
	322932 AA099732		EST cluster (not in UniGene)	1.65
	337272		CH22_FGENES.660-1	1.64
10	332694 AA262768	Hs.243901	KIAA1067 protein	1.64
	318996 Z44266		EST cluster (not in UniGene)	1.64
	315336 AW342028	Hs.256112	ESTs	1.64
	313329 AW293704	Hs.122658	ESTs	1.64
	318088 AW295409	Hs.137945	ESTs	1.64
	313835 Al538438	Hs.159087	ESTs	1.64
15	320035 AA378974	Hs.130720	ESTs; Weakly similar to CELLULAR NUCLEIC ACID BINDING PROTEIN [H.sapiens]	1.64
	309372 AW074330		EST singleton (not in UniGene) with exon hit	1.63
	324157 AW402236		EST cluster (not in UniGene)	1.63
	323929 AA354940	Hs.145958	ESTs	1.63
	302490 AA885502	Hs.187032	ESTs	1.63
20	333942		CH22_FGENES.301_8	1.63
	327469		CH.02_hs gjl5867772	1.63
	301918 AA476777		EST cluster (not in UniGene) with exon hit	1.63
	315664 Al744068	Hs.160712	ESTs	1.63
25	304405 AA282572		EST singleton (not in UniGene) with exon hit	1.63
	310624 Al341594	Hs.157522	ESTs; Moderately similar to env protein [H.sapiens]	1.63
	319250 F11623		EST cluster (not in UniGene)	1.63
	310608 Al962234	Hs.196102	ESTs	1.63
	317348 Al348076	Hs.831	3-hydroxymethyl-3-methylglutaryl-Coenzyme A lyase (hydroxymethylglutaricaciduria)	1.63
30	306513 AA989230		EST singleton (not in UniGene) with exon hit	1.63
	320807 AA086110	Hs.188536	Homo sapiens clone 24838 mRNA sequence	1.63
	303710 Al269069	Hs.250852	ESTs; Highly similar to ubiquitin hydrolyzing enzyme I [H.sapiens]	1.63
	328291		CH.07_hs gjl5868363	1.63
	304236 W93278		EST singleton (not in UniGene) with exon hit	1.63
35	317683 Al791700	Hs.127893	ESTs	1.63
	311960 AW440133	Hs.189690	ESTs	1.62
	312834 Al028309	Hs.114246	ESTs	1.62
	325326		CH.11_hs gjl5866875	1.62
	313663 Al953261	Hs.169813	ESTs	1.62
40	327526		CH.02_hs gjl6381882	1.62
	300429 AW449679	Hs.156739	ESTs; Highly similar to XG GLYCOPROTEIN PRECURSOR [H.sapiens]	1.62
	305169 AA663131		EST singleton (not in UniGene) with exon hit	1.62
	316621 Al021996	Hs.122138	ESTs	1.62
	329666		CH.14_p2 gjl6272129	1.62
45	318035 Al744130	Hs.131201	ESTs	1.62
	300492 AL031709		multiple UniGene matches	1.62
	316532 Al307229	Hs.184304	ESTs	1.62
	332048 AA496019	Hs.201591	ESTs	1.62
	307113 Al183686		EST singleton (not in UniGene) with exon hit	1.62
	319127 N49476		EST cluster (not in UniGene)	1.62
50	331155 R87650	Hs.33439	ESTs; Weakly similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]	1.61
	338220		CH22_EM:AC005500.GENSCAN.246-9	1.61
	315763 AW515270	Hs.118342	ESTs	1.61
	323571 AA984133	Hs.153260	c-Cbl-interacting protein	1.61
55	312240 R28628	Hs.203669	ESTs	1.61
	304569 AA490934		EST singleton (not in UniGene) with exon hit	1.61
	313179 Al076101	Hs.131704	ESTs	1.61
	326858		CH.20_hs gjl6552462	1.61
	317276 Al823847	Hs.129986	ESTs	1.61
60	312572 AA350125	Hs.187499	ESTs	1.61
	311932 AW451654	Hs.257482	ESTs	1.61
	302103 AA452310	Hs.26090	ESTs; Weakly similar to T20B12.1 [C.elegans]	1.61
	308413 Al636253	Hs.196511	EST	1.61
	310077 Al620617	Hs.148565	ESTs	1.61
65	337780		CH22_EM:AC000097.GENSCAN.121-2	1.61
	327796		CH.05_hs gjl5867982	1.61
	308352 Al610791		EST singleton (not in UniGene) with exon hit	1.61
	324539 Al378032	Hs.125892	ESTs	1.61
	303232 AA437414		EST cluster (not in UniGene) with exon hit	1.61
	337884		CH22_EM:AC005500.GENSCAN.54-2	1.61

	303620	AA397546	Hs.119151	ESTs	1.61
	303481	AA336839		EST cluster (not in UniGene) with exon hit	1.61
	314481	AA548589	Hs.105846	ESTs	1.61
	300327	AI908894	Hs.245893	ESTs	1.6
5	323473	AA262442		EST cluster (not in UniGene)	1.6
	326154			CH.17_hs gij5867170	1.6
	331920	AA446885	Hs.99087	ESTs; Moderately similar to ZINC FINGER PROTEIN 141 [H.sapiens]	1.6
	323827	AW406878		EST cluster (not in UniGene)	1.6
	322452	W56710		EST cluster (not in UniGene)	1.6
10	310597	AI739071	Hs.158515	ESTs	1.6
	307871	AI368665		EST singleton (not in UniGene) with exon hit	1.6
	322215	AF088005		EST cluster (not in UniGene)	1.6
	318420	AI139857	Hs.143837	ESTs	1.6
	332217	H98987	Hs.102383	EST	1.6
15	324937	M79230	Hs.192398	ESTs	1.6
	320543	AF052176	Hs.158529	Homo sapiens clone 24457 mRNA sequence	1.6
	300674	AW467388		EST cluster (not in UniGene) with exon hit	1.6
	315193	AI241331	Hs.131765	ESTs	1.6
	319713	R24204		EST cluster (not in UniGene)	1.6
20	301210	AI379982	Hs.158944	ESTs	1.6
	309365	AW072861		EST singleton (not in UniGene) with exon hit	1.6
	321403	AW451454	Hs.247568	adenylate kinase 3	1.6
	321908	AA376936	Hs.20998	ESTs	1.6
	303349	AA382661		EST cluster (not in UniGene) with exon hit	1.6
25	324338	AL138357	Hs.247514	ESTs	1.6
	310599	AW300144		EST cluster (not in UniGene)	1.6
	333193			CH22_FGENES.98_15	1.6
	336433			CH22_FGENES.825_12	1.6
	312097	AI352096	Hs.157169	ESTs	1.6
30	311445	AW204237	Hs.192703	ESTs; Weakly similar to !!!! ALU SUBFAMILY J WARNING ENTRY !!!! [H.sapiens]	1.59
	317736	AI361722	Hs.192410	ESTs	1.59
	308147	AI498991		EST singleton (not in UniGene) with exon hit	1.59
	313489	AA017492	Hs.135655	ESTs	1.59
	316289	AA902488	Hs.122952	ESTs	1.59
35	326983			CH.21_hs gij5867657	1.59
	314781	AW205298	Hs.202372	ESTs	1.59
	328397			CH.07_hs gij5868397	1.59
	331970	AA461084	Hs.187677	ESTs	1.59
	321744	N91419	Hs.12028	ESTs	1.59
40	310509	AI292181	Hs.150036	ESTs	1.59
	315921	AI147545	Hs.114172	ESTs	1.59
	322049	AI928242	Hs.144383	ESTs	1.59
	301161	AA731518		EST cluster (not in UniGene) with exon hit	1.59
45	300548	AI026836	Hs.114689	ESTs	1.59
	319142	F07366		EST cluster (not in UniGene)	1.59
	313526	AW152263	Hs.249243	ESTs	1.59
	305937	AA883238		EST singleton (not in UniGene) with exon hit	1.58
	330123			CH.19_p2 gij6671869	1.58
	327819			CH.05_hs gij5867968	1.58
50	318250	AI478814	Hs.134603	ESTs	1.58
	306760	AI034094	Hs.169476	tubulin; alpha; ubiquitous	1.58
	322358	AA220235	Hs.246836	ESTs	1.58
	317866	AI690269	Hs.201345	ESTs	1.58
	320725	AA703319	Hs.120967	ESTs	1.58
55	311332	AW292247	Hs.255052	ESTs	1.58
	334893			CH22_FGENES.452_7	1.58
	318730	AA398215		EST cluster (not in UniGene)	1.58
	315889	AW271639	Hs.221744	ESTs	1.58
60	303702	AW500748	Hs.224961	ESTs; Weakly similar to 73 kDA subunit of cleavage and polyadenylation specificity factor [H.sapiens]	1.57
	315086	AI492660	Hs.170935	ESTs	1.57
	332514	AA156499	Hs.8454	protein kinase; cAMP-dependent; regulatory; type II; alpha	1.57
	335549			CH22_FGENES.576_10	1.57
	329532			CH.10_p2 gij3983505	1.57
65	323140	AA180467		EST cluster (not in UniGene)	1.57
	313166	AI801098	Hs.151500	ESTs	1.57
	337896			CH22_EM:AC005500.GENSCAN.56-3	1.57
	330658	AA319514	Hs.211093	ESTs	1.57
	324585	AI823969	Hs.132678	ESTs	1.57

	317151	AW298195	Hs.255735	ESTs	1.57
	308818	AI819700	Hs.208231	EST	1.57
	326547			CH.19_hs gjl5867307	1.57
5	318833	H06234	Hs.24888	ESTs	1.57
	320488	R31386		EST cluster (not in UniGene)	1.57
	306929	AI124514		EST singleton (not in UniGene) with exon hit	1.57
	338083			CH22_EM:AC005500.GENSCAN.174-1	1.57
	316868	AI660898	Hs.195602	ESTs	1.57
10	310937	AI472880	Hs.170480	ESTs	1.57
	328638			CH.07_hs gjl6004473	1.57
	310074	AI651039	Hs.148559	ESTs	1.56
	327058			CH.21_hs gjl6531965	1.56
	320076	AI653733	Hs.204079	ESTs	1.56
15	322345	AF086529		EST cluster (not in UniGene)	1.56
	314731	AI745498	Hs.204579	ESTs	1.56
	318687	H49619	Hs.127301	ESTs	1.56
	303841	AI934464		EST cluster (not in UniGene) with exon hit	1.56
	302370	AJ009849	Hs.199297	Homo sapiens GNAS1 gene encoding NESP55	1.56
20	322571	AF156271		EST cluster (not in UniGene)	1.56
	318050	AI052093	Hs.133132	ESTs	1.56
	303388	AL039604		EST cluster (not in UniGene) with exon hit	1.56
	323758	AA833858		EST cluster (not in UniGene)	1.56
	328369			CH.07_hs gjl5868388	1.56
	329415			CH.Y_hs gjl5868874	1.56
25	303915	AW468839	Hs.257767	EST	1.56
	338794			CH22_EM:AC005500.GENSCAN.528-1	1.56
	303074	AA243481	Hs.127320	ESTs; Weakly similar to KIAA0346 [H.sapiens]	1.56
	318807	F08434		EST cluster (not in UniGene)	1.56
30	334287			CH22_FGENES.369_17	1.56
	311928	AW024798	Hs.233374	ESTs	1.55
	304592	AA505833	Hs.162017	EST	1.55
	300785	AA682913	Hs.247179	ESTs; Weakly similar to KIAA0319 [H.sapiens]	1.55
	304921	AA603092		EST singleton (not in UniGene) with exon hit	1.55
35	324605	AW502851	Hs.249978	ESTs	1.55
	324473	AW501163		EST cluster (not in UniGene)	1.55
	300566	H86709	Hs.21371	son of sevenless (Drosophila) homolog 1	1.55
	314165	AA761265	Hs.221281	ESTs	1.55
	302868	AA157392		EST cluster (not in UniGene) with exon hit	1.55
40	314034	AI299137	Hs.154214	ESTs	1.55
	325389			CH.12_hs gjl5866921	1.55
	331849	AA417078	Hs.193767	ESTs	1.55
	320536	AA331732	Hs.137224	ESTs	1.55
	303347	AA258033		EST cluster (not in UniGene) with exon hit	1.55
45	315769	AA744875	Hs.189413	ESTs	1.55
	317031	AA973297	Hs.126101	ESTs	1.55
	300203	AI827065	Hs.224877	ESTs	1.55
	304037	T26438		EST singleton (not in UniGene) with exon hit	1.55
	322613	AW160507		EST cluster (not in UniGene)	1.54
50	317987	AW138174	Hs.130651	ESTs	1.54
	322313	AF086386		EST cluster (not in UniGene)	1.54
	323992	AW411383	Hs.169688	ESTs	1.54
	325303			CH.11_hs gjl5866908	1.54
	312701	AI457663	Hs.128127	ESTs	1.54
55	304787	AA582678		EST singleton (not in UniGene) with exon hit	1.54
	305849	AA861571		EST singleton (not in UniGene) with exon hit	1.54
	314557	AA401367	Hs.128647	ESTs	1.54
	316507	AI381515	Hs.158381	ESTs	1.54
	315023	AA533505	Hs.185844	ESTs	1.54
60	314920	AA513406	Hs.152307	ESTs	1.54
	323097	Z44354	Hs.180950	guanine nucleotide binding protein (G protein); q polypeptide	1.54
	325043	W27919	Hs.32944	inositol polyphosphate-4-phosphatase; type I; 107kD	1.54
	307892	AI376086	Hs.158759	EST	1.54
	324573	AA491600	Hs.161942	ESTs	1.54
65	313092	AI923673	Hs.212827	ESTs	1.54
	324696	AA641092	Hs.257339	ESTs	1.54
	303019	AF098363		EST cluster (not in UniGene) with exon hit	1.54
	317158	AI459140	Hs.129109	ESTs	1.54
	309536	AW151933		EST singleton (not in UniGene) with exon hit	1.54
	301568	AI146423	Hs.146709	ESTs	1.53

5	315674	AA651923	Hs.191850	ESTs	1.53
	321861	N79341		EST cluster (not in UniGene)	1.53
	310890	AI184510	Hs.143728	ESTs	1.53
	330036			CH.17_p2 gij6042048	1.53
	316907	AA843868	Hs.190567	ESTs	1.53
10	312299	AA972712	Hs.174818	ESTs	1.53
	331128	R51361	Hs.23423	ESTs	1.53
	305177	AA663591		EST singleton (not in UniGene) with exon hit	1.53
	337685			CH22_EM:AC000097.GENSCAN.77-1	1.53
	335290			CH22_FGENES.527_3	1.53
15	308896	AI858667		EST singleton (not in UniGene) with exon hit	1.53
	307944	AI418246		EST singleton (not in UniGene) with exon hit	1.53
	300867	AW340374	Hs.121033	neural precursor cell expressed; developmentally down-regulated 1	1.53
	335320			CH22_FGENES.534_7	1.53
	329841			CH.14_p2 gij6672062	1.53
20	317916	AI565071	Hs.159983	ESTs	1.53
	332901			CH22_FGENES.36_2	1.53
	305413	AA724659		EST singleton (not in UniGene) with exon hit	1.53
	316707	AI016387	Hs.184406	ESTs	1.53
	313693	AW469180	Hs.170651	ESTs	1.53
25	316101	AA922236	Hs.221037	ESTs	1.53
	320796	AF038966	Hs.184543	secretory carrier membrane protein 1	1.53
	307451	AI248615		EST singleton (not in UniGene) with exon hit	1.53
	323648	AI679968	Hs.152060	ESTs	1.53
	331482	N27515	Hs.40296	ESTs	1.53
30	318059	AI023175	Hs.167022	ESTs	1.53
	325958			CH.16_hs gij5867142	1.53
	315736	AA664265	Hs.230213	ESTs	1.53
	314740	AW015667	Hs.119427	ESTs	1.52
	314117	AA224368	Hs.185164	ESTs	1.52
35	301646	AA313954		EST cluster (not in UniGene) with exon hit	1.52
	338752			CH22_EM:AC005500.GENSCAN.513-10	1.52
	309314	AW009312		EST singleton (not in UniGene) with exon hit	1.52
	301445	AI208364	Hs.128233	ESTs; Weakly similar to REGULATOR OF CHROMOSOME CONDENSATION [H.sapiens]	1.52
	308501	AI685263	Hs.201150	EST	1.52
40	312330	AA635305	Hs.121574	ESTs	1.52
	318040	AI018150	Hs.148781	ESTs	1.52
	336205			CH22_FGENES.719_10	1.52
	325701			CH.14_hs gij5867028	1.52
	315009	AW189460	Hs.208358	ESTs	1.52
45	303121	AW407585	Hs.27769	ESTs; Weakly similar to mCAC [M.musculus]	1.52
	309271	AI986221		EST singleton (not in UniGene) with exon hit	1.52
	328385			CH.07_hs gij5868395	1.52
	307700	AI318545		EST singleton (not in UniGene) with exon hit	1.52
	314591	AW103292	Hs.245328	ESTs	1.52
50	304484	AA432067	Hs.258373	ESTs	1.52
	304382	AA232873		EST singleton (not in UniGene) with exon hit	1.52
	304232	W52674		EST singleton (not in UniGene) with exon hit	1.52
	309853	AW298169	Hs.57553	tousled-like kinase 2	1.52
	312504	AW207346	Hs.143202	ESTs	1.52
55	313134	N63406	Hs.258697	ESTs	1.52
	330391	AF015950	Hs.115256	telomerase reverse transcriptase	1.52
	314342	AI873046	Hs.258775	ESTs	1.51
	305977	AA887293		EST singleton (not in UniGene) with exon hit	1.51
	301165	N85789	Hs.224155	ESTs; Weakly similar to PTERIN-4-ALPHA-CARBINOLAMINE DEHYDRATASE [H.sapiens]	1.51
60	300613	AI932294	Hs.249604	ESTs; Weakly similar to B-CELL LYMPHOMA 6 PROTEIN [H.sapiens]	1.51
	324124	AI554212	Hs.185664	ESTs; Weakly similar to SERINE/THREONINE-PROTEIN KINASE NRK2 [H.sapiens]	1.51
	308037	AI458207	Hs.174181	ESTs	1.51
	323909	AL043148	Hs.186257	ESTs	1.51
	315464	AW139500	Hs.116135	ESTs	1.51
65	306700	AI022056		EST singleton (not in UniGene) with exon hit	1.51
	337976			CH22_EM:AC005500.GENSCAN.107-1	1.51
	306855	AI083982		EST singleton (not in UniGene) with exon hit	1.51
	311045	AI569399	Hs.174746	ESTs	1.51
	315010	AA531082	Hs.240049	ESTs	1.51
	310205	AW025248	Hs.202445	ESTs	1.51
	310759	AW135924	Hs.224883	ESTs	1.51

	310954	AW449044	Hs.171298	ESTs	1.51
	312019	T77046	Hs.188750	ESTs	1.51
	334773			CH22_FGENES.430_5	1.51
5	332043	AA490831	Hs.125056	ESTs	1.51
	322950	AA296219		EST cluster (not in UniGene)	1.51
	337920			CH22_EM:AC005500.GENSCAN.67-3	1.51
	328993			CH.09_hs gij5868536	1.51
	309245	AI972447		EST singleton (not in UniGene) with exon hit	1.51
	312172	AI222168	Hs.191168	ESTs	1.51
10	304039	T47349		EST singleton (not in UniGene) with exon hit	1.5
	301329	AI149653	Hs.190496	ESTs	1.5
	313376	AI949246	Hs.200381	ESTs	1.5
	324248	AW504918		EST cluster (not in UniGene)	1.5
15	308771	AI809301		EST singleton (not in UniGene) with exon hit	1.5
	334935			CH22_FGENES.464_3	1.5
	319764	AA019827		EST cluster (not in UniGene)	1.5
	318519	T27135		EST cluster (not in UniGene)	1.5
	332807			CH22_FGENES.7_9	1.5
	322310	AF086376		EST cluster (not in UniGene)	1.5
20	324557	AA489166	Hs.156933	ESTs	1.5
	332118	AA609585	Hs.162689	EST	1.5
	319539	R09027		EST cluster (not in UniGene)	1.5
	313149	AW291092	Hs.201058	ESTs	1.5
25	329722			CH.14_p2 gij6065785	1.5
	323514	AA861209		EST cluster (not in UniGene)	1.5
	308078	AI472621		EST singleton (not in UniGene) with exon hit	1.5
	337965			CH22_EM:AC005500.GENSCAN.100-10	1.5
	335905			CH22_FGENES.635_13	1.5

TABLE 14A shows the accession numbers for those primekeys lacking unigeneID's for Table 14. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number
 CAT number: Gene cluster number
 Accession: Genbank accession numbers

Pkey	CAT number	Accession
322064	234514_1	BE261397 Z78343 BE176419 AA383657 N90640 AA334052 AW955761 BE536232 AA374087 AA584776
321409	197898_1	N71838 AA282003 T54072 AA761419 H92966 AI831371 AI095435 AI690247 R99331 AW964110 AA975590 AA346128
		H94196 C03864
322092	46678_1	AF085833 R69689 AW341677 AA923375 BE327566 AW630415 R69601 AW615339
321452	212379_2	AW962489 H64300 AA329527
313603	199797_1	AA284333 AW468119 AA284334 AA810992
320856	36098_1	AB040928 T94673 AI289313 AI536039 Z44366 BE141499 D60116 D61488 D59945 AA419503 R28090 R72986 H03255
		AI189112 AI912312 AW511018 AI401349 AW470144 C14624 AI335797 Z40300 AI014456 D60269 D60115 T16722 AI370673
		D60270
322139	46806_1	H53744 AF075088 H53797
321500	552826_1	BE004271 AI248023 AI022157 H71999
313733	441212_1	AA766346 AA809877 AA836116 AW469598 AW977404
322215	47002_1	AF088005 N51816 N51731
322235	47070_1	AF086106 AI193589 AW665594 N71795 AA722627 AW665373 AI300251
321632	286374_1	AW812795 AA419617 H87827 AW299775 AW382168 AW382133 BE171659 AW392392 BE171641 AA541393
313833	120893_1	AA766825 AA811180 AA085906 AI762946 AW977820
322310	47376_1	AF086376 W77804 W72689 AA837735
322313	47386_1	AF086386 W77947 W72708
322322	47434_1	AF086431 AA886756 AI557237
322331	47467_1	AF086467 W81444 W81445
322345	47537_1	W95298 AF086529 AI912190 AW294159 AI458747 W94782
322347	47545_1	AF086538 W95969 AI631911 W95835
322370	187612_1	AA330095 W25112 AA249401
321739	43998_1	AL080280 T73124 H02689 AL080281
321781	1511778_1	D78667 D78871 C18258
314570	280469_1	AA904776 AA405696 AA405962
300129	635249_1	AW028820 AI219068
322452	497108_2	AI147202 W56755 W56710
321861	1651920_1	N79341 N99082 N47551
323140	159551_1	AA180467 AA449184 AA464831 AA505048
322520	38916_1	T55958 T57205 AF147346
321914	85114_1	AA011603 N58604 N58611
322571	22297_1	NM_016102 AF156271 AA781868 AW152318 AW770403 AA909463 AA482996 AA758672
322574	39412_1	AF156548 AA639797 AI675267 AI825497 AI823355
314753	311451_1	AA463262 AA463615 AW160405 AW407583
300370	3910_2	AW136181 AA581939 AK001221 AA694538 AA424043 AI016272 AA098960 AA884473 AI356180 BE391633 AA437086
		AI277866 AA098827 AA992680 BE172624 AA424101 AA320776 AW962967 N77431 AW858960 AW858897 T85649
		AA357743 AI827817 AI905672
322601	577912_1	AI082395 W92924 BE048524 AW005302 AI084474 AI369330 AI827710 AW135506 AW298694
322613	34330_1	AW160507 NM_013367 AF191338 AA384939 AI445790 AA730309 BE397003 BE267753 AI979163 N50386 AW583671
		AW583608 BE074466 BE074479 BE074471 AW976283 AA604393 AW162122 W73648 AI823475 N75898 W73713
		AW470099 AW513236 AW025055 AW613115 AI923379 W58081 AW664525 AW196795 AI143619 AI565152 AA025406
		AA505846 AI685494 AA829964 N59156 N59163 R15442 AA826919 AI610221 AI200120 AA603279 AW150822 AI189513
		AI807122 AI016368 AI335868 AW583389 AI193892 AI956157 AI628879 AW591589 AW583446 AI955406 AW148396
		AI340255 AI867942 AA748525 AA876991 Z38516 AI874002 AI869474 N63100 AA429094 AA082443
		AW105663 AA693880 AW517398 AI768507 BE220851 AW978538 AA831489
		BE219300 BE327455 AL134620 R36741 R17996
		AL031709 AI249061 AA907658 AI420444

	316141 423880_2	AW303457 AA972713 AA724265
	323371 117336_2	N45114 N51465 BE087338 AI083551 AL135118 BE395609
	307700 30923_11	BE280998 BE254670 BE294951 BE564979 AW405364 AA069256 AA129837 AI559667 BE281405 AW410850 BE041153
5		AI254811 AW301340 AI613335 AW301411 AI609469 AI611607 AI611616 AI377623 AI335509 AI613544 BE043165 AI371663
		AI340452 AI612066 AW072890 AI254558 AI349884 AI370095 AI613383 AI611946 AI613353 AI307414 AI318229 AI612685
		AW305327 AW268924 AI370063 AI349292 BE049068 AI369098 AW274098 AI344845 AW075187 AI053401 AI345220
		BE138515 AI613386 AI583302 AW301955 AI349661 AI307432 AI054168 AI223913 AI612081 AI348942 AI334539 AI309366
		AI370098 AI252360 AW086316 AW268911 AW073482 AI379802 AI224284 AI053661 AI334538 AI309369 AI309688 AI310023
10		AI492709 AI335418 AI053999 AI366989 AW073478 AI247058 AI249584 AI305875 AI308585 AW071272 AI271487 AI340719
		AI366995 AI223673 AW271066 AI611936 AW071296 AI270796 AI254385 AI251393 AI252562 AW268236 AI254858
		AW071317 AI309102 AI609897 AW268971 AI583267 AI792484 AW075168 BE138443 AI254126 AI309822 AI310872
		AI611953 AI251054 AW276658 AI335405 AW075039 AI311768 AI612028 AW271895 AI612005 AI312240 AW271082
		AI371642 AI334879 AI310194 AI310772 AI345419 AI334675 AI223914 AI284707 AI284813 AI349140 AI254853 AI313094
15		AI310170 AI309499 AI312476 AI376484 AI335467 AI340802 AI309815 AI310168 AI611446 AI345824 BE327775 AI318545
		F17185 AW614950
	308362 792518_1	AW998989 AI613519
	307783 697809_1	AI347274 AW844024
	301161 427238_1	AA731518 AA765714
	324094 270098_1	BE395109 AW663898 AW237041 AI492154 BE046906 AI651285 AI983290 AW002590 AI201040 F32424 AA992272
20		AW271836
	309023 4737_1	AF180681 NM_015313 AA229509 AA225792 AA216413 AI888045 BE005205 AB002380 T55518 BE276097 AW380669
		BE142836 AW370976 AA479384 R96425 AI680999 AA595138 H54582 AI022709 T55440 AI041769 AA861144 AW392028
		AA479287 AA824634 AI638446 H54691 R96382 AA770352 AI640467 AW293491 AA779138 R28298 AA970562 C15590
25		R84455 AA020769 AL036394 H80566 BE548861 AA301207 AW959414 AI284253 AA043173 W52429 BE544571 R24852
		Z42603 F13120 R24340 R24326 T75305 H70110 N56255 AA334210 F11453 AW947285 H80345 AA298992 AW380931
		AI267175 Z45421 AW380981 W86113 AA663590 AA167577 BE566760 BE169166 AA449904 AA459205 N31126 W03564
		N31208 AW993277 N44765 AW605275 D61449 W68572 AA258190 D60496 AW992964 U46277 H04097 AA370360
		AW957211 AA159775 AI631243 H83367 H21671 D61077 AW392712 N21112 H98522 N45298 N83629 AI393509 AW022043
30		AA744886 AI580482 AA723286 AI422244 AI423984 D62804 AI088349 AA587890 AI144172 N33275 BE074397 H03399
		D62578 AI056639 AI829918 AA579584 AI089460 AI350124 W68573 AI580828 H98897 AI570468 H83715 W86114 AA923123
		D57446 AA043174 AW337721 AI266551 AI140017 AW022356 D79855 D79650 D79393 D60495 AA788666 AA693443
		AW516977 W60139 AI628156 AW473223 AI608892 AA159670 AW440366 AI421529 T50751 AI174374 AA912234 AA724248
		AW780400 AA907218 H80514 D57452 AA863419 AA552618 D29614 R44556 T16452 R44935 Z41132 D29188 H69692
35		AI250176 AI078860 AA370359 AW183108 H74200 AA258183 F10723 C00323 R86148 AA860570 AW130073 AL079946
		AA410327 AA532614 AA234500 AI151507 AA410288 AW969839 AA483232 AI383200 AA236540 AI807672 H73441
	323473 193878_1	AA262442 AA768862 AA262443
	315639 392767_1	AA827650 AA827652 AW629526 BE044585 AW974451 AA761439 AA648505 AA765803
	322878 117013_1	AA081820 AA082191 AA079811
	301239 457668_1	AA807558 AA827117 AW629567
40	301256 16720_1	NM_016603 AF251038 AI124624 AA776579 AW298470 AI304868 AW082724 AI348442 BE218336 N20641 AI018013
		AW858832 AW978157 AA815187 AA932948 AF157316 AI444958 W00848 W02935 AI434933 N26335 AA428681 AW371059
		AI651612 AW134937 AW968911 AA488815 AI157523 W48766 AW936954 AW936941 AW579205 AW936886 AW936889
		N74541 AW936953 AW578421 AW604352 AW367088 AW849258 AW849453 AW371606 AI554921 W49785 H99814
45		AA805957 AA904606 AW206696 BE169229 AA333951 AA190704 AW936944 AA463219 AA430306 AW805704 N48503
		BE222307 AI638612 BE550045 AI805304 AI690987 AA776841 H12690 AW183731 AI380760 AI636261 AA812641
		AW592656 AI686132 AA843424 H99220 AW084996 AW128879 AI800871 AA610135 AA191524 AI150076 AI474530
		AA748461 N29013 AA746372 N56606
	300611 337193_1	N75450 AA877636 AW137945 W05248 AA514763 AW972399 AI758397 AW195051
	324157 247225_2	AW402931 BE393099
50	323509 967739_1	AL036947 T93676 T85475
	323514 197787_1	AA641735 AA281881 AA861209 AA934756 AA835887 AA641795 AA748822 AW295703
	300674 466093_1	AW467388 AA826954
	322932 39838_1	AF168711 AA099732 BE019157 AI380212 BE298159 AA249097 AA305112 AW962349 AW962353 AW401801 BE292961
55		AI439469 AA442919 AI630537 AA724473 AI814288 AW966815 AI376871 AI860202 AI683132 AA099733 AW627633
		AI754022 BE206347 AW183349 AI378222 BE178926 AI473282 W52944 AW752469 AW966817
	323591 209807_1	AA301270 AA301379 AA301366
	322950 10774_1	R85652 AA114024 AA296219 AA375304 AW963796 AW885952 AW020969 AA114025 AI804930 BE350971 AI765355
		AW317067 AW974763 H85930 AW172600 AI310231 AW612019 D62908 D62864 AA652738 AI674617 AI494064 AW138666
60	322957 29014_1	AI147620 AI147629 AW611793 AI668922 AI971005 AI864742 AA174171
		AK001701 AA134337 AA356202 BE163251 AW875175 AW875181 AW875177 BE163389 AK000741 AA247755 AA120819
		AW868040 AA309118 AW962348 AA471267 AW996843 AK001452 BE005344 BE617899 AA186588 AA120820 AW363311
		AA648105 N71529 BE168417 AW673900 AI858160 AA134338 AA659697 N22162 AI335437 AI311237 AI343171 AI336661
		AW268074 AW274348 AA935005 AW576295 AW262626 AW593153 AA730055 AA662650 AA782687 AW894855 AI933533
65		AW193002 AW899448 AW890142 AW812670 AA085664 AA334191 BE178085 BE180553 AA389680 AA984772 AA442527
		W26560 BE384359 AA847210 AW304931 AI669606 AA085613 AW197240 AI632828 AA581646 AW129348 AI017643
		AW089030 D20893 AI382955 AI557148 AW499979
	324231 975669_1	W60827 AL079968 AL047234
	324248 977901_1	AW504918 N55410 AL118584 AW839266
	323691 221757_1	AA317561 AI793000 AW235111 AI793178 AA767397 AI263113 AA719462

	315858	406384_1	AA737345 AA682286 AI799378
	301431	569736_1	R05385 AI061251
	324303	233842_1	AL118754 AA333202 H38001
5	324330	300543_1	AA884766 AW974271 AA592975 AA447312
	300815	41537_2	BE152396 BE152395 AA287515 BE001834 AA286678 AW406477
	324349	1154015_1	AW501470 AW502931 AW499500
	323715	225129_1	AA322155 AA326396 AA326538
	309314	23273_-3	AW009312
10	323758	229624_1	AA833858 AW978090 AA327679 AA810436
	309375	127_1	AF286598 AW075342 AB028994 AL043713 AW378914 AA340650 N57166 AW956914 R17961 AA336481 BE393734
			AW977867 AW294638 AA927857 AA961627 AW303969 AW894416 AA812119 AA912758 AA424355 AA490582 W30941
			AA476693 AA131029 AA127777 AL043714 AA496984 T51117 AA127722 AA594012 AI492876 N76483 AW119061 BE464926
			AW303419 AI972370 AI768172 AI826550 AI435432 AI379516 AA778421 AI276089 AA424521 N59361 AA723153 AA723176
			AI867487 AA090677 AI827221 AI351027 W02732 AI810729 AA142848 AI082110 N59379 N29744 AI283747 AI148665
15			AW779845 AI382967 F34319 AI369934 AI282438 AW183449 AA863467 AA813469 AI092645 AI870701 AA863119
	325031	266373_2	T65475 R07576 T17017 F08143 Z43546
	325045	1534945_1	T08845 Z43538 F06691
	324473	38795_1	BE560824 BE513941 AW238907 AA580852 AW501176 BE241846 AW501163 AW751433 AW501340 BE241715 AI910774
	323827	235506_1	AW406878 AW966560 AW966151 AW966496 AA336174 AA335376 AA335537
20	302270	1734192_1	R56151 W91936
	301618	10967_5	T52761 T52760
	301646	42154_1	AJ277841 AI630669 AI804370 Z41939 AW751251 AA299456 Z44739 AW860471 Z30158 AW105391 H56997 W84688
			AA491201 W84636 AA706815 AI131055 AA483636 AI005075 AW340034 AI332372 AW118195 AI338932 AI191968
			AA693932 AI189982 AI193225 AA884163 AA594562 W37747 AA249754 AA746131 AI916540 AI832188 AW946555
25			AA833838 Z40564 AA861563 F01447 AA887937 AI933559 AW973250 AA566018 AA313954
	323923	249295_1	AA354146 AI184230 AA643525
	324580	328264_1	AA492588 AA492498 AA492571
	316774	463723_1	AA814859 AA814857 AI582623
	309577	6483_6	AW902251 AW168753
30	302345	29533_1	X12830 NM_000565 AW503691 X58298 S72848 AA193347 AW503481 AW177946 AW178192 AW178188 AA285233
			AA410577 AA193465 AW177939 AW365459 BE221693
	302358	1064753_1	AW207734 D60164 D81150 D81078 D61356 AW996804
	324614	215437_1	AW503101 AA309184 N56323 R70998
35	324661	385257_1	AW504161 AW503601 AW505509
	324685	41003_1	AF226667 AA207032 AA100804 AA121287 AA488316 AI808218 AW419048 AI911097 AW132123 AA502311 AW089948
			AA100952 AI075431 AW083432 AI990554 BE466029 F28643 AF086422 W79581 AW439007 F37179 W79780 AW439035
			AA731381 AW750380 AA251012 AW589846 AA730238 AA329792 AW087255 AA220982 AA082469 AA877260 AA232380
			BE298910
	324692	351987_1	AA557952 AA677593 AA618150
40	316893	473541_1	AW979189 AA837332 AA856946 AA876935
	303027	21796_1	AF111178 NM_005708 AF105267 AW590040 AI979280 AA001322 BE146329 AA702430 AA702429 AA694221 AI206348
			AI206285 AW770197 AA923032 AI379586 AW701165 AW594643 AA001909 AW002368
	324715	290035_2	AI739168 AA426249 AI199636 AW505198 AW977291 AA824583 AA883419 AA724079 AI015524 AI377728 AW293682
			AI928140 AA731438 AI092404 AI085630 AA731340
45	324771	385085_1	AA631739 AA768584 AW134477
	324783	389615_1	AA640770 AI683112 AA913009
	303114	37417_1	AF090948 AI064898 AI111182
	303124	21112_1	AB018257 BE148640 AA081832 AK001915 AF150217 AF161350 AI219174 AW074664 D60040 AA346065 H28750
50			AW151783 BE613360 BE612628 BE502031 AW183790 AA992580 AA505815 AI310432 AI678015 AW592679 AA879181
			AA806708 AI744110 H24681 C16064 D62900 AI285033 AA346064 AI865123 AW467798 BE221231 AL120676 N89877
			AI928370 AI358387 AA748486 AV647478 AV647460 AA312313 AI279340 AW505099
	302552	82290_1	AA005122 H49792
	301918	316229_1	AA476777 T86049
55	303232	20474_1	AA437414 AA131479 AA086182 AB037775 AW161063 AW514393 AA332331 AW136197 BE150789 AA425533 AA249605
			N88308 AI016201 BE004662 AA291027 R57587 AA424277 AA476391 W07532 T97036 AA218898 AW162629 R57770
			W01278 W90204 W90156 AL119197 R84513 AA280103 AA334994 AW965504 AA460868 AA447470 AW138594 W38898
			W90028 AI078353 W90078 AA699696 N35523 AA704225 AA035059 AW134892 AA115140 AI142854 H90084 AA826342
			AA460694 N46339 AA425344 N56953 AA035569 AI761083 AI658696 AI524818 AI338965 AW069249 AW299871 BE464061
60			AI189720 AW340682 AI423380 AI275122 H17532 N80735 AA826343 AI039694 BE328398 AI192947 AW271286 AI623122
			AI922902 AW293087 N22141 AA730657 AW316610 N26473 F06663 Z43610 H14783 R59761 H11540 AI265915 AI681773
			AI091748 BE220636 AW841861 AI702181 AI468447 AA907544 AI273941 AW244034 R37769 AA446663 T96929 BE045884
			AA476341 H89994 H29043 AW051211 N49522 AA306977
	302696	33570_1	AK000738 AA347452 AW961713 H70832 AI750643 AA362887 AW955588 W44974 AA279599 AW298762 AA452666
			AA443355 AI337273 AA446931 AI752977 AA661554 W42674 AI292172 R41163 AA621381 AI244157
65	302697	43219_1	AJ001409 AJ001410
	309917	57485_2	AW340014 AW866993 AV651649
	303347	192210_1	AA258033 AA459485
	303349	193138_1	AA382661 AW958642 AA259088
	310599	690880_1	AW300144 AI338491 AI798381 BE220076

5	303388	969232_1	AL039604 AL039497
	302761	45074_1	AW250553 L07876 Z36843 R30693 AI190097 AW965317
	318455	606341_1	AI148763 AI903763 AI903753 AI903762 AI903800 AI903801
	317850	363835_1	AI681545 AI951714 AI570397 AW873588 AA836396 AI359986 AI499790 AA773477 AI951615 T07547 AW304709 AF114041
	303431	32082_1	BE176629 Z44580 T30422 T32690 AW953065 H10602
10			NM_000539 AA019013 AA019367 AA056154 H38735 AA057003 AA021051 H38102 AA015774 AA059291 AA019439 H84843
			H83375 AA019914 AA017288 R84449 W26519 H38258 AA018736 H84147 AA018577 AA059353 U49742 H38767 AA318341
			AA317553 H86646 H91989 AA317398 AA317378 W29024 W23034 T27877 AW950059 AA017195 R84262 AA057177
			H89941 AA019904 H84662 AA015775 AA019368 AA020976 H37900 C20733 H38682 H85197 AA018578 AA017252
			AA019440 AA059059 H38651 H84148 AA018560 W25754 C20752 AA317915 AW952115 AA317369 AA019845 R85402
15	319127	1653640_1	AA019492 AA017196 AA056093 AA056094 AA058836 AA056155 W25957 W23027 AA056159 W23043 W21890 W28951
	303480	232749_1	AA317978 W26459 AA317265
	303481	31534_1	N49476 Z45911 R21061
			AA331906 AA332484
			AK001952 AA336839 AW249271 BE247287 AF182002 BE613472 AW962673 AA332235 AW849937 AW849814 H49893
20			AA477148 AW968944 AF182003 AW007897 BE246145 W76100 AI480141 AW410205 AA609339 AI209111 AW000979
			AA330280 AW961554 W72865 H49894 AA514317 AA620407 AA504522 AW472833 AA716609 AW129282 AA347351
			AA628378 AW589860 AI636696 AA464632 AA464533 AW874189 AA757076 AA479654 AW517910 AW292357 AW872638
			AW262288 AI910666 AW513749 AW238771 AA215797 BE387073
	303487	20890_1	BE143533 AW850432 AK000042 AA333666 AA385314 AW966616 AW793068 AW793414 AA361103 AW390841 AA040095
25			AW385058 AW799162 AI383115 AI990745 AI653703 BE503693 AW150758 AI949919 AW190450 AW512348 AI625970
			AW501057 N52954 AI281378 AI401710 AI648409 AW002659 AI687639 AI093943 R33960 AA040062 AI926267 AI240425
			AI520911 AI093428 R52943
	303488	36085_1	AI040372 AB040915 W40569 BE158910 BE158914 D63226 AW025860 AW583088 AA334307 AA210942 AW753212
			AW805322 AA362635 BE158911 AW891225 AW994862 AA805451 R28541 AA229347 N48266 AI377788 R28682 R36122
30			AA811941 AI240742 AI632001 T99965 W01976 AW891205 AW891177 T97433 C15571 AA346850 AA504293 W07500
			AI694503 AA489216 AA327725 AW959917 AA694146 N68514 AI076285 AW016246 T07783 AA642400 AA716133 AA805332
			R00312 AA705021 AW498605 AW891723 AW891906 AA808025 N29039 N74897 W60393 AA810184 AI627460 AW057516
			AA807436 AA760966 AI359295 N78642 N20662 AA830300 W81705 AA832258 AW891718 AI811796 AW515523 Z41735
			AA449978 AW891714 AI684539 AW891896 AW071701 AI890916 AI924994 AI039743 AA888524 AA244214 AI015736
35			AI270105 AI865077
	303494	236389_1	F30712 F35665 AW263888 AI904014 AI904018 AA336927 AA336502
	319142	164820_1	H08370 Z46168 F07366 AA193168 AA193138
	302868	12593_1	AK000290 AI476034 AA465309 BE148761 AW303607 AW958665 AW469635 AI819365 AI243857 AW469326 AA157110
			AA278626 AA496257 AA306656 F29732 AA831859 AA312210 AA564476 AA579065 AA769522 AA740386 AI205635
40			AA491643 AA810400 AA417708 AI567332 AA157392 N53817 AA374229
	318518	1205335_1	R68545 T27119 R25687 AW750672
	318519	434741_1	H13364 T27135 R61679 AA746905
	304168	72494_-10	H77679
	302948	21445_1	AB038995 NM_016530 AK001111 AA465635 AW968716 U66624 AA885459 AA703019 AI040266 AI018689 AI692886
45			AI125372 AI376796 AI192040 N58161 AL133607 AW503673 AW505479 AA362265 AJ404671
	319250	244351_1	F11623 H17552 AA347728
	318644	17700_1	BE311816 AK000916 AW868037 AW868039 AF228527 AI752482 AW868041 AA077049 AI201537 W55873 AA206019
			AA077918 AW968729 AI978828 AW139620 AI093053 AW204025 AI418805 AA598926 AA586345 AA045669 BE314455
			AA045668
50	318674	204968_1	W01166 AW996900 BE184300 Z44887 T34535 R51495 AW886575 AA295490 AA295162 AA295163 AW937125 T56951
	304232	20640_2	BE386106 W52674
	303685	8088_1	AW500106 BE241915 AW503971 NM_016542 AB040057 AA313812 AK000556 W16504 AI822088 AA259107 AA191319
			BE085957 AA309584 BE122687 AW952435 T84469 BE088194 BE088132 AA328562 BE092674 AA263102 T39634
			AW992380 R79391 R24392 H03060 AW675066 AI299952 AW020325 D25953 N75199 AA361425 AW612302 AW236333
55			AW673897 AW953686 N22323 AA649166 AI377099 H03061 AI660072 AW276405 AA809779 AI803430 AW297484
			AW510384 AA814816 AA371522 D63035 AA953567 R79392 R24282 AA876831 AW297542 AI699023 AA992652 AI041436
	318704	799152_1	AI631602 AW589676 Z28684 Z24981
	318730	275116_1	Z32887 BE349923 AA398215 AA399231
	303714	1155758_1	AW501336 AW501337
60	304387	183612_1	AA236027 BE003275
	304398	10169_1	AA195509 BE394661 AV660757 AA489161 BE165972 AW503705 AA262785 AF123320 Z78357 NM_014171 AF161488
			AA248971 BE568575 AA461410 AA165108 AI637731 H75454 AA372934 AW339334 BE568754 BE564697 BE567299
			AI681606 BE537269 AW197204 AA290890 AI189393 AW292463 AW470227 F27399 AW611942 BE566888 AW301701
			AI675761 AI628429 AA164711 AI797753 AI656879 AI912690 AI675277 AI695099 AI094095 AW014158 BE091059 AI201748
65			AW236961 AI038003 AI083606 AA401606 AI079405 AI073516 AI655537 AA401475 AI814532 AI079862 AI093789 AI422084
			AI216476 AI392760 AA926998 AA781782 Z25198 AI086377 AI185511 AI185539 Z28843 AI223792 AI379563 AA706253
			AI433798 AI921885 H75455 AW025269 AI224100 AI083611 AI225057 AW196334 AI572254 AA761628 AI472801 AA283784
	303751	468554_1	AA830149 AW978407 M85983 AW503637
	319401	1323199_1	W00973 N56457 AW992226 T84921 R01342
	319402	1003489_1	R86913 R86901 H25352 R01370 H43764 AW044451 W21298
	318807	1536467_1	F08434 Z42573 H28810
	319478	765461_1	AI524124 R06841 R06842
	318872	1534581_1	Z43108 F06295 R13085

	318885	94880_2	AA742999 Z43272 AA345258 AW956677 AA031942
	303841	79133_1	W19657 BE616760 BE259848 BE382680 BE615587 AI934464 AA322745 T07155 AW961174 AA307302 Z41888 AA621992
			AA188400 AW770608 AI147458 AI148408 AI696291 AA972591
5	303889	1777183_1	T19204 T36109 T36107
	319539	63198_1	R09027 AA344892 AA329574 AW955648 AW978708 AI567804 AI378935 AW014657 AI804134 R08922 N92947 BE546788
	318905	1536408_1	F08365 Z43395 R54298
	320187	396254_1	T99949 AA654769 AA664550 AW975264
	318996	65715_1	Z44266 H06384 AV655948
10	319635	163534_1	R17531 AW960899 AA338366 AW673294 BE047729 BE047722 AA330746 AW841797 H05030 AI142105 R12654
	319699	747196_1	AI458682 H24240 R14537 R18426 AW867082
	319713	1699356_1	R24204 R15712 T84695
	319761	75324_2	AW630974 BE005208 R84237 AA724997 AA334867 AW955777 R18816
	319764	88596_1	AA019827 R18947 H46852
15	319808	7069_3	T58960 AA609180 AA621130 AI927236 AA431075
	321040	193331_1	AA261830 AW967855 H26953 AA262478
	320409	43709_1	AA226869 AA296516 AW959753 AA186390 AL359619 AA356195 AA148427 R22748 AI033624 BE548853 H95327
20			AW579751 BE561649 AA397533 BE617136 AA236444 T89946 AA247450 N55777 W38725 AI743846 AI808406 AA922229
			AI051464 W04713 R11251 W19656 AI042319 AA489276 AI224533 H95274 AW269958 T89311 AI890088 AI862754
			AI830968 AI669336 AI589780 AA534557 AW273839 AI338155 AI126632 N83542 BE046048 AA807028 AA848107
			AW167978 AA976930 AA148428 AI289304 AI524262 AI625961 AA773469 AI222288 AI280054 AI242371 AA227222
			AA973329 AA296517 AA829436 AA234526 AI149769 AI567865 AA936939 AI590681 AW469308 AI689531 AA486419
			AI422051 AI057252 AA626941 AI475352 AW247913 AI222370 AA670122 AW198034 AA486418 AI363794 AA380739
	319881	1585983_1	H51299 H44619 H46391 R86024 H51892 T72744
25	320488	368456_1	AI817336 R32883 AA595590 AI743065 R31386
	321121	1545647_1	W23285 H42714 F25381 F37215
	321205	81249_1	AA002047 N72537 H54142 H81580
	321253	375160_1	AA610649 AI699484 H59558
	314043	155125_1	AA827082 AA732246 AA167611 AA830741
30	320630	17685_2	AA199847 AA410224 R53323 AW936567 AW936569 AW936568 AW936571
	313435	443527_1	AA769123 AA831715 AW977666 W92553
	313443	82292_1	AA005125 W95019 W93335 AA249037
	313472	82811_1	AA007374 AA007466 AI816886
	321348	41762_1	Z49979 D61703 U30168
35	314138	179960_1	AA740616 AA654854 AA229923
	320712	57156_2	R66867 R65678 R82673 W73128 R83101
	321383	41924_1	AW968556 AJ238555 AW968731 AJ002574 AA459446 H70260 AW977557 AA767351 AW268572 AA810719 AI698677
			AI300460 AA907450 AA649224 T07415 AI536896 BE018515 AI279865 BE047421
			AW368634 AI702169 AI245179 AW368646 BE545574 AA249018 AW368633 N27553
40	312996	187327_1	AA989230
	306513		AA991705
	306557		AA994530
	306598		AI000320
	306620		AI000929
45	306700		AI022056
	308078		AI472621
	306813		AI066544
	306830		AI075803
	306855		AI083982
50	329722	c14_p2	
	329728	c14_p2	
	306890		AI092235
	308100		AI475949
	308147		AI498991
55	306929		AI124514
	308352		AI610791
	308383		AI624497
	308521		AI689808
	308561		AI701559
60	308617		AI738720
	308771		AI809301
	308828		AI824829
	308896		AI858667
	303019	41850_1	AF098363 AF098365
65	303084	44211_1	AF174008 AF174027 AF174106
	305092	AA642912	
	305169		AA663131
	305177		AA663591
	305235		AA670480
	305413		AA724659

	305849	AA861571
	305854	AA862733
	307113	AI183686
5	307130	AI185234
	305937	AA883238
	305977	AA887293
	307451	AI248615
	307513	AI274307
10	307848	AI364186
	307871	AI368665
	307881	AI370434
	307932	AJ230822
	307944	AI418246
15	307954	AI419692
	307965	AI421641
	309245	AI972447
	309271	AI986221
	309365	AW072861
20	309372	AW074330
	309435	AW090537
	309506	AW137700
	309536	AW151933
	309709	AW242630
25	325417 c12_hs	
	325450 c12_hs	
	325452 c12_hs	
	309815	AW292760
	309839	AW296076
30	309849	AW297444
	309906	AW339340
	302705 31765_1	U09060 U09061
	304037	T26438
	304039	T47349
35	304236	W93278
	304257	AA053294
	304382	AA232873
	304405	AA282572
	304561	AA489792
40	304569	AA490934
	304787	AA582678
	304921	AA603092
	327819 c_5_hs	
	304968	AA614308
	306382	AA968967
45	331263 47479_1	AW780192 AA015718 W02571
	332252 1663967_1	N63882 T91174

TABLE 14B shows the genomic positioning for those primekeys lacking unigene ID's and accession numbers in Table 14. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique number corresponding to an Eos probeset
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers
 Strand: Indicates DNA strand from which exons were predicted.
 Nt_position: Indicates nucleotide positions of predicted exons.

	Pkey	Ref	Strand	Nt_position
15	332807	Dunham, I. et.al.	Plus	297686-297808
	332808	Dunham, I. et.al.	Plus	298277-298360
	332812	Dunham, I. et.al.	Plus	309688-310561
	332901	Dunham, I. et.al.	Plus	1841954-1842090
20	333149	Dunham, I. et.al.	Plus	3574317-3574413
	333916	Dunham, I. et.al.	Plus	8298994-8299169
	334026	Dunham, I. et.al.	Plus	9196549-9196681
	334061	Dunham, I. et.al.	Plus	9686941-9687077
	334073	Dunham, I. et.al.	Plus	9792201-9792374
25	334150	Dunham, I. et.al.	Plus	10529221-10529854
	334379	Dunham, I. et.al.	Plus	13908356-13908467
	334719	Dunham, I. et.al.	Plus	15778859-15779026
	334773	Dunham, I. et.al.	Plus	16235169-16235328
	334893	Dunham, I. et.al.	Plus	19302753-19302881
30	334935	Dunham, I. et.al.	Plus	20108247-20108373
	335146	Dunham, I. et.al.	Plus	21491292-21491457
	335320	Dunham, I. et.al.	Plus	22542132-22542246
	335568	Dunham, I. et.al.	Plus	24935021-24935655
	335586	Dunham, I. et.al.	Plus	24990333-24990497
35	335601	Dunham, I. et.al.	Plus	25044923-25045157
	336036	Dunham, I. et.al.	Plus	29019796-29019877
	336123	Dunham, I. et.al.	Plus	30051089-30051186
	336268	Dunham, I. et.al.	Plus	31997555-31998040
	337173	Dunham, I. et.al.	Plus	23624127-23624224
40	337460	Dunham, I. et.al.	Plus	32536159-32536395
	337685	Dunham, I. et.al.	Plus	3547161-3547245
	337736	Dunham, I. et.al.	Plus	3850500-3850643
	337780	Dunham, I. et.al.	Plus	4113793-4113990
	337965	Dunham, I. et.al.	Plus	7034267-7034392
45	337976	Dunham, I. et.al.	Plus	7166011-7166119
	338030	Dunham, I. et.al.	Plus	8072708-8072827
	338112	Dunham, I. et.al.	Plus	10391398-10391600
	338165	Dunham, I. et.al.	Plus	12205719-12205875
	338178	Dunham, I. et.al.	Plus	12800037-12800181
50	338427	Dunham, I. et.al.	Plus	19685043-19685354
	338506	Dunham, I. et.al.	Plus	21221871-21221953
	338794	Dunham, I. et.al.	Plus	27114697-27114763
	338910	Dunham, I. et.al.	Plus	28795375-28795551
	339047	Dunham, I. et.al.	Plus	30760793-30760968
55	332864	Dunham, I. et.al.	Minus	1390386-1390296
	332933	Dunham, I. et.al.	Minus	2035790-2035681
	333193	Dunham, I. et.al.	Minus	3832993-3832494
	333712	Dunham, I. et.al.	Minus	7286177-7286073
	333940	Dunham, I. et.al.	Minus	8523830-8523671
60	333942	Dunham, I. et.al.	Minus	8552629-8552330
	334287	Dunham, I. et.al.	Minus	13294116-13293871
	334387	Dunham, I. et.al.	Minus	13946021-13945781
	334487	Dunham, I. et.al.	Minus	14432191-14432132
	334913	Dunham, I. et.al.	Minus	19463909-19463815
65	335109	Dunham, I. et.al.	Minus	21325792-21325667
	335250	Dunham, I. et.al.	Minus	21952922-21952826

	335288	Dunham, I. et.al.	Minus	22304275-22303770
	335290	Dunham, I. et.al.	Minus	22309950-22309891
	335549	Dunham, I. et.al.	Minus	24666203-24666128
5	335862	Dunham, I. et.al.	Minus	26690300-26690125
	335864	Dunham, I. et.al.	Minus	26694537-26694382
	335905	Dunham, I. et.al.	Minus	26988888-26988719
	336205	Dunham, I. et.al.	Minus	30477456-30477311
	336276	Dunham, I. et.al.	Minus	32093320-32093181
10	336433	Dunham, I. et.al.	Minus	34067540-34067425
	336605	Dunham, I. et.al.	Minus	15616509-15616358
	336616	Dunham, I. et.al.	Minus	26021027-26020848
	336679	Dunham, I. et.al.	Minus	2035790-2035681
	337043	Dunham, I. et.al.	Minus	17407330-17407251
	337272	Dunham, I. et.al.	Minus	28241476-28241307
15	337357	Dunham, I. et.al.	Minus	30906179-30906109
	337393	Dunham, I. et.al.	Minus	31471747-31471569
	337497	Dunham, I. et.al.	Minus	33371317-33371258
	337646	Dunham, I. et.al.	Minus	2648689-2648632
	337920	Dunham, I. et.al.	Minus	6051648-6051510
20	338083	Dunham, I. et.al.	Minus	9318438-9318301
	338220	Dunham, I. et.al.	Minus	14166440-14166104
	338752	Dunham, I. et.al.	Minus	26421374-26421135
	338763	Dunham, I. et.al.	Minus	26628148-26628009
	338983	Dunham, I. et.al.	Minus	29908865-29908702
25	339209	Dunham, I. et.al.	Minus	32492953-32492593
	325240	5866848	Minus	32301-32650
	329532	3983505	Plus	42937-43014
	329522	3983507	Minus	35265-35458
30	329519	3983510	Plus	18407-18597
	329511	3983514	Plus	20965-21325
	325326	5866875	Plus	47726-48024
	325303	5866908	Minus	73556-73630
	325389	5866921	Plus	239672-239759
35	325417	5866925	Minus	110635-110745
	325450	5866941	Minus	435379-435552
	325452	5866941	Minus	704103-704202
	325498	5866967	Plus	173372-173930
	325587	6682462	Plus	126724-126967
40	325602	5866994	Plus	79122-79251
	325701	5867028	Minus	72936-73046
	325780	6381953	Plus	63634-63873
	329722	6065785	Minus	112713-112992
	329728	6065785	Minus	207544-207741
	329666	6272129	Plus	98307-98446
45	329815	6624888	Minus	68431-68720
	329841	6672062	Minus	40181-40331
	325824	5867048	Minus	42450-42833
	325866	5867076	Minus	94358-94628
	325902	5867101	Minus	127729-127842
50	325958	5867142	Plus	53437-53550
	326014	5867160	Minus	10358-10447
	329941	6165199	Minus	34319-34411
	330002	6623963	Plus	46097-46158
55	326154	5867170	Minus	7103-7179
	326023	5867245	Plus	171799-171896
	326278	5867269	Plus	75250-75903
	330036	6042048	Plus	117120-117216
	326547	5867307	Minus	623677-623870
60	326495	5867423	Plus	11843-11930
	326507	5867435	Minus	13038-13111
	326505	5867435	Minus	8818-8949
	326506	5867435	Minus	9368-9509
	326530	5867441	Minus	303000-303122
	326508	6682496	Plus	78904-79112
65	330120	6671864	Minus	127553-127656
	330123	6671869	Minus	35311-35406
	326858	6552462	Minus	69337-69670
	326983	5867657	Minus	16023-16581
	327014	5867664	Plus	1017630-1017788

	326930	6456782	Plus	606950-607705
	326920	6456782	Minus	42425-42519
	327058	6531965	Plus	2384268-2384835
5	327061	6531965	Minus	3486389-3486673
	327075	6531965	Plus	4041318-4041431
	327120	6531970	Minus	6-1088
	330126	6093735	Plus	82458-82623
	327157	5866841	Minus	4408-4746
10	327183	5867442	Plus	84317-84531
	327192	5867445	Minus	194652-194764
	327288	5867481	Plus	48583-48773
	327469	5867772	Plus	145549-145708
	327489	6004459	Minus	57796-58015
15	327526	6381882	Minus	97010-97123
	327574	5867818	Plus	68767-69126
	327665	5867839	Plus	141736-141900
	327752	5867949	Plus	93721-94421
	327819	5867968	Minus	92202-92717
20	327796	5867982	Plus	85267-85405
	330260	6671884	Plus	45203-45269
	330282	6671910	Plus	3982-4114
	328078	5868008	Plus	72807-72865
	328121	5868031	Plus	153782-153850
25	328190	5868077	Plus	21082-21165
	328227	5868105	Minus	21082-21242
	327871	5868131	Minus	88889-89221
	328018	5902482	Minus	542547-543133
	328624	5868246	Minus	120666-120836
30	328744	5868290	Plus	138639-138722
	328799	5868316	Minus	80771-80923
	328291	5868363	Minus	144244-144434
	328329	5868375	Plus	191709-192239
	328369	5868388	Plus	75371-75583
35	328385	5868395	Plus	369952-370155
	328397	5868397	Plus	344967-345063
	328412	5868405	Plus	86427-86519
	328538	5868485	Plus	3814-4243
	328656	6004473	Plus	792616-792729
40	328638	6004473	Plus	294618-294903
	328903	5868514	Plus	23625-24468
	328960	6456775	Plus	38547-38837
	330320	5932415	Minus	54458-54697
	328993	5868536	Plus	49160-50084
45	329081	5868602	Plus	93368-93510
	329089	5868614	Plus	25805-26923
	329109	5868626	Plus	102168-102273
	329192	5868716	Plus	166936-167020
	329218	5868726	Minus	71408-71707
50	329224	5868728	Plus	27422-27664
	329246	5868732	Minus	250541-250792
	329415	5868874	Plus	1011438-1011818
	329454	5868887	Plus	51342-51593

TABLE 15: 169 GENES WITH SEQUENCE INFORMATION DEPICTED IN TABLE 16

Table 15 depicts UnigeneID, UnigeneTitle, Primekey, Predicted Cellular Localization, and Exemplar Accession for all of the sequences in Table 16. The information in Table 15 is linked by EosCode to Table 16.

Pkey: Unique Eos probeset identifier number
 ExAccn: Exemplar Accession number, Genbank accession number
 UnigeneID: Unigene number
 Unigene Title: Unigene gene title
 EosCode: Internal Eos name
 Localization: Predicted cellular localization of gene product

	Pkey	ExAccn	UnigeneID	Unigene Title	EosCode	Localization
	100394	D84276	Hs.66052	CD38 antigen (p45)	PBC1	plasma membrane
	100452	D87742	Hs.241552	KIAA0268 protein	PAB7	not determined
	101249	L33881	Hs.1904	protein kinase C, ι ota	OAA1	cytoplasmic
20	101485	M24736		selectin E (endothelial adhesion molecul	ACC5	plasma membrane
	101514	M28214	Hs.123072	RAB3B, member RAS oncogene family	PFJ2	cytoplasmic
	101851	M94250	Hs.82045	midkine (neurite growth-promoting factor	LBH9	secreted
	102398	U42359		gb:Human N33 protein form 1 (N33) gene,	PDG3	
	102522	U53347	Hs.183556	solute carrier family 1 (neutral amino a	PFJ4	plasma membrane
25	102669	U71207	Hs.29279	eyes absent (Drosophila) homolog 2	LEM9	cytoplasmic
	103119	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta	LBG2	plasma membrane
	103709	AA037316	Hs.13804	hypothetical protein dJ462O23.2	PDO6	
	104080	AA402971	Hs.57771	kallikrein 11	PBA6	secreted
	104144	AA447439	Hs.183390	hypothetical protein FLJ13590	PDM3	
30	104691	AA011176	Hs.37744	Homo sapiens beta-1 adrenergic receptor	PAV1	plasma membrane
	105370	AA236476	Hs.22791	transmembrane protein with EGF-like and	PDM9	plasma membrane
	106149	AA248881	Hs.256301	hypothetical protein MGC13170	PDO8	
	106579	AA456135	Hs.23023	ESTs	PAA4	plasma membrane
	107102	AA609723	Hs.30652	KIAA1344 protein	PAA3	not determined
35	107217	D51095		DKFZP586E1621 protein	PDG8	
	108153	AA054237	Hs.40808	ESTs	PBF1	plasma membrane
	109014	AA156790	Hs.262036	ESTs, Weakly similar to Z223_HUMAN ZINC	PDG7	
	109112	AA169379	Hs.257924	hypothetical protein FLJ13782	BCU4	not determined
	109890	H04649	Hs.20843	Homo sapiens cDNA FLJ11245 fis, clone PL	PDG4	
40	110151	H18836	Hs.31608	hypothetical protein FLJ20041	PAV9	plasma membrane
	112971	T17185	Hs.83883	transmembrane, prostate androgen induced		CHA1 not determined
	113021	T23855	Hs.129836	KIAA1028 protein	PDO3	
	114908	AA236545	Hs.54973	cadherin-like protein VR20	PFJ6	plasma membrane
	114965	AA250737	Hs.72472	ESTs	BCY2	mitochondrial
45	116393	AA599463		hypothetical protein MGC2648	PDV3	secreted
	116416	AA609219	Hs.39982	ESTs	OAB6	
	117698	N41002	Hs.45107	ESTs	PDT9	ER
	117984	N51919	Hs.106778	ATPase, Ca++ transporting, type 2C, memb		PAJ5 not determined
	118985	N94303	Hs.55028	ESTs, Weakly similar to I54374 gene NF2	PDM8	
50	119018	N95796	Hs.278695	Homo sapiens protein mRNA, complete cds		-PAB2 plasma membrane
	119126	R45175	Hs.117183	ESTs	PBF8	
	120992	AA398246	Hs.97594	KIAA1210 protein	PDG5	
	121710	AA419011		prostate androgen-regulated transcript 1	PDV5	
	121913	AA428062		ESTs; protease inhibitor 15 (PI15)	BCU7	vesicular
55	122041	AA431407	Hs.98732	Homo sapiens Chromosome 16 BAC clone CIT		PAZ1 not determined
	122593	AA453310	Hs.128749	alpha-methylacyl-CoA racemase	PDO1	
	123209	AA489711	Hs.203270	ESTs, Weakly similar to ALU1_HUMAN ALU S		PAA2 plasma membrane
	124526	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci	PAV4	plasma membrane
	126399	AA128075		transmembrane, prostate androgen induced	PDY4	
60	126645	AI167942	Hs.61635	six transmembrane epithelial antigen of	PAA5	plasma membrane
	126966	R38438	Hs.182575	solute carrier family 15 (H+/peptide tra	PDO5	plasma membrane
	127537	AA569531	Hs.162859	ESTs	PAA6	not determined
	128790	AA291725	Hs.105700	secreted frizzled-related protein 4	BCX2	secreted
	129109	AA491295	Hs.108708	calcium/calmodulin-dependent protein kin	PFJ7	
65	129184	W26769	Hs.109201	CGI-86 protein	PAV6	vesicular
	129389	AA621604		spondin 2, extracellular matrix protein	CJA5	not determined

	129404	AA172056	ESTs	PAB4	
	129534	R73640	Hs.11260 hypothetical protein FLJ11264	PAJ3	secreted
	130760	AA128997	Hs.18953 phosphodiesterase 9A	PEE6	nuclear
	131425	AA219134	Hs.26691 ESTs	PBA7	
5	132964	AA031360	ESTs	PAA7	plasma membrane
	132967	AA032221	Hs.61635 six transmembrane epithelial antigen of	PM17	plasma membrane
	133179	U81599	Hs.66731 homeo box B13	PFJ5	nuclear
	133330	U42360	Hs.71119 Putative prostate cancer tumor suppresso	PDM1	plasma membrane
	133520	X74331	Hs.74519 primase, polypeptide 2A (58kD)	PDM2	
10	133724	U07919	Hs.75746 aldehyde dehydrogenase 1 family, member		PDT1 mitochondrial
	133724	U07919	Hs.75746 aldehyde dehydrogenase 1 family, member		PDT1 mitochondrial
	133944	AA045870	Hs.7780 Homo sapiens mRNA; cDNA DKFZp564A072 (fr		PAB9 cytoplasmic
	134110	U41060	Hs.79136 LIV-1 protein, estrogen regulated	BCR4	plasma membrane
	301805	AI800004	Hs.142846 hypothetical protein	PEU4	nuclear
15	302005	AI869666	Hs.123119 MAD (mothers against decapentaplegic, Dr	PBJ6	cytoplasmic
	302881	AA508353	Hs.105314 relaxin 1 (H1)	PBH3	secreted
	303506	AA340605	Hs.105887 ESTs, Weakly similar to Homolog of rat Z	PEG4	
	303699	D30891	Hs.19525 hypothetical protein FLJ22794	PBM4	not determined
	303753	AW503733	Hs.9414 KIAA1488 protein	PBY3	not determined
20	308050	AI460004	Hs.31608 hypothetical protein FLJ20041	PEU5	plasma membrane
	310382	AI734009	Hs.127699 KIAA1603 protein	PCQ8	
	310431	AI420227	Hs.149358 ESTs, Weakly similar to A46010 X-linked	PBH1	plasma membrane
	310573	AW292180	Hs.156142 ESTs	PEN3	plasma membrane
	310598	AI338013	Hs.140546 ESTs	PCW3	
25	310816	AI973051	Hs.224965 ESTs	PET5	
	311596	AI682088	Hs.79375 holocarboxylase synthetase (biotin-[prop	PBH8	
	313676	AA861697	Hs.120591 ESTs	PBY2	
	314121	AI732100	Hs.187619 ESTs	PBY1	
	314691	AW207206	Hs.136319 ESTs	BFF8	not determined
30	314785	AI538226	Hs.32976 guanine nucleotide binding protein 4	CBO7	cytoplasmic
	314907	AI672225	Hs.222886 ESTs, Weakly similar to TRHY_HUMAN TRICH		PBM2not determined
	315051	AW292425	ESTs	PBM9	
	315052	AA876910	Hs.134427 ESTs	PBJ7	plasma membrane
	316442	AA760894	Hs.153023 ESTs	PBJ9	
35	317548	AI654187	Hs.195704 ESTs	PBQ6	
	317869	AW295184	Hs.129142 deoxyribonuclease II beta	PBQ7	
	318524	AW291511	Hs.159066 hypothetical protein FLJ10188	PBJ1	cytoplasmic
	319191	AF071538	prostate epithelium-specific Ets transcr	PEN1	
	319763	AA460775	Hs.6295 ESTs, Weakly similar to T17248 hypotheti	PEO7	
40	320324	AF071202	Hs.139336 ATP-binding cassette, sub-family C (CFTR	PBH5	plasma membrane
	320561	NM_006953	Hs.159330 uroplakin 3	PEL9	plasma membrane
	320796	AF038966	Hs.31218 secretory carrier membrane protein 1	PBY4	not determined
	321441	AW297633	Hs.118498 Homo sapiens LUCA-15 protein mRNA, splic		PBY8 not determined
	322303	W07459	Hs.157601 ESTs	CBF9	secreted
45	322782	AA056060	Hs.202577 Homo sapiens cDNA FLJ12166 fis, clone MA		PBQ1 not determined
	322818	AW043782	Hs.293616 ESTs	PCQ7	plasma membrane
	323226	AF055019	Hs.21906 Homo sapiens clone 24670 mRNA sequence		PCI2 not determined
	323287	AA639902	Hs.104215 ESTs, Moderately similar to SPCN_HUMAN S		PBJ5
	324295	AI146686	Hs.143691 ESTs	PBQ9	not determined
50	324430	AA464018	Hs.184598 Homo sapiens cDNA: FLJ23241 fis, clone C		PBY6 not determined
	324603	AW016378	Hs.292934 ESTs	PBM3	
	324617	AA508552	Hs.195839 ESTs, Weakly similar to I38022 hypotheti	PBH4	cytoplasmic
	324626	AI685464	gb:tt88f04.x1 NCI_CGAP_Pr28 Homo sapiens	PCW6	
	324658	AI694767	Hs.129179 Homo sapiens cDNA FLJ13581 fis, clone PL		PBJ4 plasma membrane
55	324718	AI557019	Hs.116467 small nuclear protein PRAC	CBK1	nuclear
	330211			PBJ2	not determined
	330546	U31382	Hs.299867 guanine nucleotide binding protein 4	PEW1	cytoplasmic
	330762	AA449677	Hs.15251 hypothetical protein	PBM1	not determined
	330790	T48536	Hs.122764 TMPRSS2, transmembrane protease, serine		PEL3 plasma membrane
60	330892	AA149579	Hs.91202 ESTs	PBQ4	plasma membrane
	331099	R36671	Hs.14846 Homo sapiens mRNA; cDNA DKFZp564D016 (fr		PCQ1cytoplasmic
	331490	N32912	Hs.291039 ESTs	PCI4	nuclear
	331889	AA431407	Hs.98802 ESTs, Moderately similar to T14342 NSD1	PBH7	not determined
	332247	N58172	gb:za21f09.s1 Soares fetal liver spleen	PBQ5	nuclear
65	332396	AA340504	gb:hw31a09.x1 NCI_CGAP_Kid11 Homo sapien		PBJ8 not determined
	332697	T94885	transgelin 2	PBQ8	secreted
	332798			PBH2	nuclear
	334447			PBY9	not determined
	338255			PBY7	not determined

	401424			PFG2	mitochondrial
	407122	H20276	Hs.31742	ESTs	PEW7
	408430	S79876	Hs.44926	dipterydylpeptidase IV (CD26, adenosine	PEZ3
5	408826	AF216077	Hs.48376	Homo sapiens clone HB-2 mRNA sequence	PEY1
	409262	AK000631	Hs.52256	hypothetical protein FLJ20624	PFG1
	409361	NM_005982	Hs.54416	sine oculis homeobox (Drosophila) homolo	PEW3
	411096	U80034	Hs.68583	mitochondrial intermediate peptidase	PEZ9
	413125	BE244589	Hs.75207	glyoxalase I	PFJ3
	413623	AA825721	Hs.246973	ESTs	OBH6
10	414422	AA147224	Hs.337232	Homeo box A13	PFC6
	415263	AA948033	Hs.130853	ESTs	PEZ5
	417153	X57010	Hs.81343	*collagen, type II, alpha 1 (primary ost	PFJ1
	418601	AA279490	Hs.86368	calmegin	PFA1
	418848	AI820961	Hs.193465	ESTs	PEY4
15	418882	NM_004996	Hs.89433	ATP-binding cassette, sub-family C (CFTR	OBH2
	419839	U24577	Hs.93304	*phospholipase A2, group VII (platelet-a	PFH9
	421887	AW161450	Hs.109201	CGI-86 protein	PFH2
	422083	NM_001141	Hs.111256	*arachidonate 15-lipoxygenase, second ty	PFH5
	424565	AW102723	Hs.75295	guanylate cyclase 1, soluble, alpha 3	PFA3
20	425071	NM_013989	Hs.154424	*deiodinase, iodothyronine, type II"	PFH6
	425710	AF030880		solute carrier family, member 4	PFD4
	427958	AA418000	Hs.98280	potassium intermediate/small conductance	PFH1
	428819	AL135623	Hs.193914	KIAA0575 gene product	PFD6
	429900	AA460421	Hs.30875	ESTs	PEZ7
25	429918	AW873986	Hs.119383	ESTs	PEY5
	430226	BE245562	Hs.2551	adrenergic, beta-2-, receptor, surface	PEZ4
	431217	NM_013427	Hs.250830	Rho GTPase activating protein 6	PFG6
	431716	D89053	Hs.268012	fatty-acid-Coenzyme A ligase, long-chain	PEZ1
	431992	NM_002742	Hs.2891	protein kinase C, mu	PFH4
30	432189	AA527941		gb:nh30c04.s1 NCI_CGAP_Pr3 Homo sapiens	PFA2
	432244	AI669973	Hs.200574	ESTs	PEW8
	432437	W07088	Hs.293685	ESTs	PFG3
	432966	AA650114	Hs.325198	ESTs	PEY3
35	439176	AI446444	Hs.190394	ESTs, Weakly similar to B28096 line-1 pr	PEW5
	440260	AI972867	Hs.7130	copine IV	PEW6
	440901	AA909358	Hs.128612	ESTs	PFC8
	445424	AB028945		cortactin SH3 domain-binding protein	PEZ6
	446320	AF126245	Hs.14791	*acyl-Coenzyme A dehydrogenase family, m	PFH7
	447210	AF035269		phosphatidylserine-specific phospholipas	PFH8
40	449156	AF103907	Hs.171353	prostate cancer antigen 3, non-coding DD	PEZ8
	449625	NM_014253		odz (odd Oz/ten-m, Drosophila) homolog 1	PEZ2
	449650	AF055575	Hs.23838	calcium channel, voltage-dependent, L ty	PFD2
	451939	U80456	Hs.27311	single-minded (Drosophila) homolog 2	PFJ8
	451982	F13036	Hs.27373	Homo sapiens mRNA; cDNA DKFZp564O1763 (f	PFG9
45	452039	AI922988		ESTs	PFD8
	452340	NM_002202	Hs.505	ISL1 transcription factor, LIM/homeodoma	PFG4
	452784	BE463857	Hs.151258	hypothetical protein FLJ21062	PFC5
	452946	X95425	Hs.31092	EphA5	PFH3

TABLE 15A shows the accession numbers for those primekeys lacking a unigeneID in Table 15. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number
 CAT number: Gene cluster number
 Accession: Genbank accession numbers

Pkey	CAT number	Accession
------	------------	-----------

116393	131543_1	AI972402 AI634409 AI523716 AI799749 W44518 AI424438 AI688513 AI971048 AI686324 AW013854 AA588483 AA528111 AI627428 AI582200 AI669296 AI826926 AI620526 AI669958 AI972458 AI924500 AA512903 W44517 AA335363 AW238997 BE300165 BE250665 AA284195 AA523420 W52834 AI471970 AI952824 AW003820 AW009463 AA669796 AA114966 AI653342 AA115038 AI342150 AI092100 AI968211 W51994 AI804005 AI201420 AI123210 AI738405 AI674964 AI970341 AW027500 AI493316 AI333193 AI139353 AA599463 AI656163 AI804200 AI365321 AI990213 AI657011 AA650025 AI968810 AI341978 AA599839 AW592602 AA644289 AI468578 AI565265 AI565228 BE221535 AW973052
101485	18113_1	AA296520 AL021940 M30640 NM_000450 M24736 M61894 AL047443 H39560 AI694691 AA916787 AI214796 AA939085 AI150616 AA412553 AA412545 AI051015 T27654 AA694430
126399	17331_1	AA088767 AF224278 AA128075 AL035541 AA027926 AI761441 AI972096 AW071693 AI742327 AI377498 AI804815 AI640802 AI885001 AI921394 AA595115 N71820 AI921217 AW007283 AI467828 AI369306 AA917446 AI493698 AA088701 AA126899 AI936228 AW204238 AI039567 AI925027 BE138909 AW452945 AW135998 AA310984 AA027860 AW073519 AI537597 AA953976 AI521341 AW273569 AW050740 AA536113 AA559064 AI474392 AW135709 AA535181 AW572959 AA570597 AI905464 AI677810 AI587642 AW975102 AA424310 AA482527 N64192 AA658276 AW889117 AA486591 AW889172 AI381990 AI381991 AI673419 AI990950 AA487031 AI272934 AI150565 AA229168 AW316722 AI142707 BE222396 AA614168 AA122026 AW338227 AA632457 AI968726 AW369662 AA512956 AA541675 AA451748 AI250993 BE146418 AA122025
132964	94346_1	AI362575 AI805082 AW263421 AI432462 AA135870 AA031360 AA031604 AA298475 AA298464
129389	21074_1	NM_012445 AB027466 BE407510 BE047605 AA047125 AW084003 AA149494 AA149490 AA292528 AA570505 AA526186 AW006250 AW007762 AI341557 AI799666 AI972710 AI377966 AI962810 AI084783 AI458032 AI190971 AW148913 AA372354 AW970032 AW007426 AA650188 AI123203 AI122890 AI280975 W73595 W73495 AI863238 AA374109 AA603986 AW149089 AW957523 AI307748 AI921067 AI336463 F24537 AI380460 AI367500 AI189309 AI814701 AI766921 AW572106 AA037024 AW072576 AA578293 AI288103 AA235464 AW450642 AA574230 AW294024 AI589229 AI580733 AW512227 AA877009 AI660255 AW188597 AA558228 AI572782 AA658397 AI274628 AI866359 AA864573 AI264439 AA621604 AW515493 AW243333 Z39737 AI567038 AA573997 AA573559 AW236431 AI652870 AI684973 AA034505 AA047126
129404	156454_1	AI267700 AI720344 AA191424 AI023543 AI469633 AA172056 AW958465 AA172236 AW953397 AA355086
107217	9836_1	AL080235 AA031750 D81382 AI480231 AI095947 AI560953 BE010721 AI870290 AA374945 AA125792 D51527 D51556 AI685541 D51559 AW117286 AA195741 AI675138 AW593439 AI201885 T30590 AW952100 D51095 AA523864 W70043 AA987586 AI421515 AI205532 AA127069 AI337367 D51595 AI453785 AW075677 AW088359 C14287 C14284
121710	19266_1	AF163474 NM_016590 AF163475 AI761105 AI770098 AA410580 AA411616 AI590343 AI739050 AL050198 AI862645 AA419104 AA513809 AA333032 AI816915 AW139625 AA640889 AI311391 AI627693 AW135514 AA419011 AI269149 AI245259 AI970008 AI970017 AW139445 AA569503 AI761072 AI766179 AI759995 AI300776 AI870129 AW150770 AA226501 AA226220
121913	291015_1	AI249368 AI742316 AA428062 AA442089 AI864189 BE349478 AI803475 AI584049 BE552085 AI088609 AI264197 AI886144 AI129474 AI307145 BE181300 AW058403 AI696838 AW748598 AA442196 AI216428
102398	entrez_U42359U42359	
315051	347217_1	AW292425 BE467167 AI702953 BE550961 BE222309 AI299348 AI693336 AA541708
324626	336411_1	AI685464 AW971336 AA513587 AA525142
319191	16065_1	NM_012391 AF071538 AB031549 AI685592 AI745526 AA662204 AW130657 AA662164 AW971121 AI668916 AA513274 AI991223 AI979170 AW298436 AA639821 AI859010 AW513942 AI687669 AA662521 AA548598 AI345056 AI305374 BE043418 AI432856 AI334840 AI379796 AI492693 AI307915 BE042082 AI307834 AI307858 AI309488 BE042210 AI435670 AI371605 AI862491 AI284563 AI306872 AI255044 AI254601 AI251236 AI473073 AI473042 AI432760 AI435664 AI336826 AI289365 AI369096 AI862274 AI334871 AI349863 AI250405 AI377617 AI309895 AI313017 AI862291 AI311936 AI378718 AI305722 AI306769 AI308888 AI334565 AI862296 AI344230 AI435685 AI344087 AI378696 AI311209 AI435775 AI310611 AI311154 AI432289 AI431561 AI492681 AI432867 AI335288 AI492796 AI432769 AI310299 AI432273 AI379820 AI275319 AI435753 AI609441 AI432767 AI369100 AI311420 AI439974 AI247157 AI334677 AI270910 AI224320 AI305608 AI334489 AI377152 AI350012 AI370086 AI335053 AI306781 AI306750 AI334849 AI334874 AI340380 AI307876 AI305974 AI305972 AI311521 AI334872 AI862509 AI311498 AI335051 AI289684 AI310859 AI311862 AI862483 AI492775 AI307906 AI492708 AI289693 AI340373 AI307910 AI311359 AI435653 AI334865 AI311492 AI492809 AI492690 AI431576 AI862268 AI311879 AI308435 AI492792 AI862512 AI275321 AI431564 AI431564 AI307885 AI307926 AI435692 AI435778 AI310182 AI308894 AI492707 AI492713 AI308560 AI307829 AI343234 AI580598 AW472796 AI340918 AI310243 AI309368 AI307920 AI289665

5 AI306777 AW086318 AW086292 AW086378 AI310027 AI275293 AI369082 AI340900 AI306749 AI371558 AW086287 BE043803
 AI306793 AI306272 AI287948 AI270917 AI284816 AI336813 AI284546 AI308044 AI275290 AI270872 AI306795 AI289687 AI223570
 AI305303 AI289677 AI287742 AI275284 AI306812 AI336701 AI371554 AI378719 AI344988 AI223631 AI335141 AI343222 AI284568
 AI305357 AI275270 AI345932 AI436549 AI307925 AI311502 AI344238 AI343182 AI308508 AI305988 AI270790 AI379792 AI305647
 AI305410 AI432251 AI436517 AI343227 AI305534 AI340387 AI271043 AI305499 AI271046 AI305962 AI289465 AI305378 AI289725
 AI310848 AI305848 AI289362 AI252964 AI307049 AI310831 AI306993 AI306796 AI224659 AI305969 AI349855 AI306164 AI306948
 AI284676 AI309155 AI343202 AI432785 AI306815 AI369081 AI270885 AI289699 AI435704 AI309647 AI305716 AI311281 AI287927
 AI472995 AI340423 AI270958 AI307069 AI305364 AI270807 AI275306 AI311890 AI275263 AI432750 AI289371 AI432861 AI255113
 AI305709 AI473008 AI311168 AI309711 AI377164 AI271201 AI289560 AI309710 AI306195 AI311201 AI287741 AI271066 AI432876
 AI275281 AI379795 AI472972 AI311967 AI306826 AI305465 AI270792 AI473019 AI305340 AI270922 AI305995 AI305462 AI254144
 AI270969 AI473012 AI305390 AI275278 AI223644 AI289692 AI250318 AI305372 AI289691 AI250521 AI306283 AI306814 AI307933
 AI473160 AI432903 AI223720 AI254979 AI334862 AI306926 AI289541 AI432248 AI435722 AI435698 AI432859 AI310683 AI473175
 AI335144 AI289467 AI436489 AI306928 AI473033 AI305763 AI307868 AI307882 AI348959 AI435736 AI432857 AI432896 AI435735
 AI432283 AI473086 AI432863 AI473081 AI432825 AI307840 AI473164 AI432885 AI473166 AI472982 AI435734 AI473060 AI473171
 AI432279 AI432882 AI334670 AI436512 AI432827 AI432852 AI473051 AI473077 AI435697 AI271509 AI492781 AI472983 AI473018
 AI432897 AI473043 AI432871 AI436536 AI473157 AI349715 AI432777 AI473016 AI473158 AI340369 AI307941 AI432773 AI377146
 AI492791 AI270950 AI305342 AI284604 AI306269 AI284811 AI270811 AI289347 AI334869 AI334852 AI311759 AI250382 AI309520
 AI289550 AI305721 AI340870 AI270901 AI308575 AI307904 AI340715 AI270941 AI309808 AI246867 AI473014 AI307039 AI289360
 AI473069 AI492786 AI344013 AI305876 AI436510 AI340742 AI473028 AI307891 BE041871 BE041268 BE042340 BE041946
 20 BE041783 AI306173 AI201948 AI926972 AI275769
 338255 CH22_6856FG_LINK_EM:AC00
 330211 c_5_p2
 332798 CH22_14FG_6_5_LINK_C4G1.G
 334447 CH22_1746FG_387_7_LINK_EM
 25 332247 372969_1 AA669097 AA513815 AA026798 AA676526 AA704429 AA704269 AW118292 AA579216 N58172
 332396 20265_1 AW579842 BE156562 BE156690 BE156489 BE081033 AK001559 BE149402 M85387 AW367811 AW367798 R17370 AI908947
 AA382932 R58449 H18732 AA371231 AW962899 AA713530 AW892946 R53463 H11063 AW068542 Z40761 BE176212 BE176155
 W23952 W92188 AW374883 AA303497 AW954769 AA036808 BE168063 AW382073 AW382085 AL041475 H80748 AI078161
 BE463983 AI805213 AI761264 W94885 N94502 AI623772 AI419532 AI810302 AI634190 AW002516 AW150777 AI352312 AI367474
 30 AW204807 AI675502 AI337026 AW134715 BE328451 AI123157 AI560020 AI300745 AI608631 AI248873 AA742484 AW051635
 H18646 AI245045 AA507111 AI640510 AI925594 AA115747 AA143035 AA151106
 332697 13699_1 X51405 NM_001873 T11322 AL118886 BE328175 AW136009 BE467445 AW470313 AA774852 BE504139 AW501046 AA082792
 AW389231 AA370044 R36841 AA371457 C04813 R25791 R25556 AW895854 AW903819 AW895671 AW895677 BE159723
 35 AW895664 AW895597 AW895595 AW895665 AW888518 AI903724 F06081 F08503 AL119462 AW895730 AW888516 R26511
 R26489 AA334126 AA327626 N85713 AW895998 AA223622 F05468 AA370749 W05590 M78202 AA371073 AW498607 R15017
 T16991 AA001282 AA001138 AA551566 AA330159 AI922855 AA383512 AA029603 D82246 D82171 T94933 H56545 AA348060
 AA176888 R96764 AW451817 AA385766 AA452618 AI690057 AA988822 BE549928 AA150901 W57992 AW899925 C05281
 40 AA932042 AA370980 AW962877 W04741 AA369982 AW385948 AA922466 N75882 AI422070 AI361256 AI680224 D57122 T94885
 R53266 R46713 T19071 AW796277 AA325333 F04719 F02334 AA358146 AA626597 AA358304 AW028099 AL119570 D57290
 D58273 D57796 N48555 AI361969 AA329457 D57225 AW024046 AA992606 AW022118 AW021538 AA935845 H89870 H56546
 AW961219 AA453239 AW837541 N45521 BE218029 AA318877 AA327740 AW961809 T92139 D53216 D52365 D53363 D53312
 D53116 AI547267 AA679935 AW026552 AW026418 AW190507 AI927710 AW244108 D50948 AW054991 AW021063 AW022511
 45 AA493436 AI365636 BE464751 AW149384 AA102442 AW771368 AI818251 AI126368 D51049 AI421542 AI559467 AW079779
 AW021048 AW023969 AW044214 AI458264 AA027274 AI620254 AW028917 BE219511 AA326242 N67561 AI971273 AA878328
 D57131 AA770662 AI309299 AI796767 AA613338 W58076 AI566287 AI445573 AI880260 AA001919 AW339259 AI492610 AI492611
 R97692 AI301425 AA722603 D58361 AI350323 AA973926 AI431263 AA516126 AA865467 AI925177 N39443 AA001943 AI299371
 AI082412 AA665090 AA583433 H89871 AA977231 AI362219 AI056096 AI270446 N67524 N22103 AW614224 AA744054 AW243622
 50 AI613188 AI929173 AI350243 AI362138 AA744004 AA176661 D56787 AI955625 AI393109 AI094769 AI479728 AI423107 AI955617
 AI034036 AI582196 AW264534 AI418961 AA570761 AI343538 AA650341 AA992503 AA770004 AL039666 AI862675 AW190335
 AA610274 AW418627 BE467472 D56786 T28749 AI217610 AI359556 T23523 AL040189 AA846222 AA651636 D51280 AI888986
 AI521167 AI340177 AW612815 AI625285 AA621607 AA177059 AA229768 AA829788 AI749682 AW190631 N75299 AA230089
 AI915632 BE069542 AA890020 AA528397 AA995390 BE503860 AA570812 AW339396 AI197986 AI203725 AI282379 AA670375
 AA461513 F01728 AW243599 C00856 N75567 R95995 AA150932 R95961 AA648060 AA933800 AA927073 AA101126 AA864190
 T93566 BE167472
 55 425710 25529_1 AF030880 NM_000441 AC002467 AA385554 H23053 AW891838 AI139968 AA653057 AI695233
 432189 342819_1 AA527941 AI810608 AI620190 AA635266
 445424 6391_1 AB028945 T77648 F13328 AL157605 Z46212 AA304736 F11855 T66098 T30174 AW954164 AW176301 AW748243 AA456428
 AI369958 AA938565 AW959613 Z42008 AA994779 AI683909 F11019 F10926 AI769597 AI752550 T65015 AI884314 AA643954
 Z41838 AW020147 AI038822 AW571822 AA299781 AA894928 AF131790 BE005411 AI902476 AW082695 AA464384 R42750
 60 AW902301 AA464273 R05837 Z38294 H41098 AI134507 M86079
 447210 7119_1 AF035269 AF035268 NM_015900 T96213 U37591 AA156832 AA299371 AI084325 H95977 AI765967 BE221465 AA156726 AI969563
 AW024539 AI436791 AI949451 AA843093 AI452756 AA824232 AI306667 T96131 AW207447 AW243556 AW957032 AI084332
 H95978 U30998
 65 449625 8113_1 NM_014253 AF100772 BE088769 AL022718 BE161779 AW863569 BE161640 AL039060 BE168542 AW296554 AA323193 AA235370
 AW779760 N48674 AI375997 R45432 D59344 AI203107 F07491 R35360 R25094 AI913631 AI498402 T61382 AI016320 N45526
 T61415 AA331486
 452039 89513_1 AI922988 H05475 AA021608 AW169947 AA913750 Z41614 AW800012

TABLE 15B shows the genomic positioning for those primekeys lacking unigene ID's and accession numbers in Table 15. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique number corresponding to an Eos probeset
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
 Strand: Indicates DNA strand from which exons were predicted.
 Nt_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	Nt_position
334447	Dunham, I. et al.	Plus	14308764-14308824
332798	Dunham, I. et al.	Minus	232147-231974
338255	Dunham, I. et al.	Minus	15242294-15242231
330211	6013592	Plus	59158-59215
401424	8176894	Plus	24223-24428

TABLE 11 AND SEQUENCE LISTING

SEQ ID NO:1 BCU4 DNA SEQUENCE

Nucleic Acid Accession #: NM_024915

Coding sequence: 13-1890 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | | |
 5 ATTGGATCAA ACATGTCACA AGAGTCGGAC AATAATAAAA GACTAGTGGC CTTAGTGCCC 60
 10 ATGCCCACTG ACCCTCCATT CAATACCCGA AGAGCCTACA CCAGTGAGGA TGAAGCCTGG 120
 AAGTCATACT TGGAGAATCC CTTGACAGCA GCCACCAAGG CCATGATGAT CATTAAATGGT 180
 GATGAGGACA GTGCTGTCTG CCTCGGCCCTG CTCTATGACT ACTACAAGGT TCCTCGAGAC 240
 AAGAGGCTGC TGTCTGTAAG CAAAGCAAAGT GACAGCCAAG AAGACCAGGA GAAAAGAAAC 300
 15 TGCCTTGGCA CCAGTGAAGC CCAGAGTAAT TTGAGTGGAG GAGAAAACCG AGTGCAAGTC 360
 CTAAAGACTG TTCCAGTGAA CCTTCCCTA AATCAAGATC ACCTGGAGAA TTCCAAGCGG 420
 GAACAGTACA GCATCAGCTT CCCCAGAGAGC TCTGCCATCA TCCCGGTGTC GGGAAATCAG 480
 GTGGTGAAAG CTGAAGATTT CACACCAGTT TTCATGGCCC CACCTGTGCA CTATCCCCCG 540
 GGAGATGGGG AAGAGCAACG AGTGGTTATC TTTGAACAGA CTCAGTATGA CGTGCCCTCG 600
 20 CTGGCCACCC ACAGCGCCTA TCTCAAAGAC GACCAGCGCA GCACTCCGGA CAGCACATAC 660
 AGCGAGAGCT TCAGGAGAGC AGCCACAGAG AAATTTCCGA GTGCTTCAGT TGGGGCTGAG 720
 GAGTACATGT ATGATCAGAC ATCAAGTGGC ACATTTCACT ACACCTGGA AGCCACCAA 780
 TCTCTCCGTC AGAAGCAGGG GGAGGGCCCC ATGACCTACC TCAACAAAGG ACAGTTCTAT 840
 GCCATAACAC TCAGCGAGAG CGGAGACAAC AAATGCTTCC GACACCCCAT CAGCAAAGTC 900
 25 AGGAGTGTGG TGATGGTGGT CTTCAAGTGA GACAAAAACA GAGATGAACA GCTCAAATAC 960
 TGGAAATACT GGCACCTCTG GCAGCATACG GCGAAGCAGA GGGTCCTTGA CATTGCCGAT 1020
 TACAAGGAGA GCTTTAATAC GATTGGAAAC ATTGAAGAGA TTGCATATAA TGCTGTTTCC 1080
 TTTACCTGGG ACGTGAATGA AGAGGGGAAG ATTTTCATCA CCGTGAATTG CTTGAGCACA 1140
 GATTTCCTCT CCAAAAAAGG GGTGAAAGGA CTTCCTTTGA TGATTTCAGT TGACACATAC 1200
 30 AGTTATAACA ATCGTAGCAA TAAACCCATT CATAGAGCTT ATTGCCAGAT CAAGGTCTTC 1260
 TGTGACAAAG GAGCAGAAAG AAAAAATCCG GATGAAGAGC AGAAGCAGAA CAGGAAGAAC 1320
 GGGAAAGGCC AGGCTCCCA AACTCAATGC AACAGCTCCT CTGATGGGAA GTTGGCTGCC 1380
 ATACCTTTAC AGAAGAAGAG TGACATCACC TACTTCAAAA CCATGCCTGA TCTCCACTCA 1440
 CAGCCAGTTC TCTTCATACC TGATGTTTAC TTTGCAAAAC TGCAGAGGAC CGGACAGGTG 1500
 35 TATTACAACA CGGATGATGA ACGAGAAGGT GGCAGTGTCC TGTGTTAAACG GATGTTCCCG 1560
 CCCATGGAAG AGGAGTTTGG TCCGGTGCCT TCAAAGCAGA TGAAAGAAGA AGGGACAAAG 1620
 CGAGTGTCTT TGTACGTGAG GAAGGAGACT GACGATGTGT TCGATGCATT GATGTTGAAG 1680
 TCTCCACAG TGATGGGCTT GATGGAAGCG ATATCTGAGA AATATGGGCT GCCCGTGGAG 1740
 AAGATAGCAA AGCTTTACAA GAAAAGCAAA AAAGGCATCT TGGTGAACAT GGATGACAAC 1800
 40 ATCATCGAGC ACTACTCGAA CGAGGACACC TTCATCTCA ACATGGAGAG CATGGTGGAG 1860
 GGCTTCAAGG TCACGTTCAT GGAAATCTAG CCTGGGTTT GGCATCCGCT TTGGCTGGAG 1920
 CTCTCAGTGC GTTCTCCCT GAGAGAGACA GAAGCCCCAG CCCCAGAACC TGGAGACCCA 1980
 TCTCCCCCAT CTCACAACG CTGTTACAAG ACCGTGCTGG GGAGTGGGGC AAGGGACAGG 2040
 CCCCACAGTC GGTGTGCTTG GCCATCCAC TGGCACCTAC CACGGAGCCG AAGCCTGAGC 2100
 CCCCAGGAA GGTGCCTTAG GCCTGTTGGA TTCCTATTTA TTGCCACCT TTCTCTGGAG 2160
 45 CCCAGGTCCA GGCCCCCAG GACTCTGCAG GTCACGCTA GCTCCAGATG AGACCGTCCA 2220
 CGGTTCCTCC TCAAGAGAA ACACCTACCT CGAACAGCCT AAAAAATTCC CATCCCTTCT 2280
 TTCTCACCCC TCCATATCTA TATCTCCCGA GTGGCTGGAC AAAATGAGCT ACGTCTGGGT 2340
 GCAGTAGTTA TGGGTGGGCG AAGAGGTGGA TGCCCACTTT CTGGTCAGAC ACCTTTAGGT 2400
 50 TGCTCTGGGG AAGGCTGTCT TGCTAAATAC CTCCAGGTTT CCCAGCAAGT GGCCACCAGG 2460
 CCTGTACAG GAAGACATTC AGTCACCGTG TAATTAGTAA CACAGAAAGT CTGCCTGTCT 2520
 GCATTGTACA TGTGTTTAT AATATTGTAA TAATATATTT TACCTGTGGT ATGTGGGCAT 2580
 GTTTACTGCC ACTGGCTAG AGGAGACACA GACCTGGAGA CCGTTTAAAT GGGGGTTTTT 2640
 GCCTCTGTGC CTGTTCAAGA GACTTGCAGG GCTAGGTAGA GGGCCTTTGG GATGTTAAGG 2700
 55 TGATGCGAGC TGATGCCAAG ATGGACTCTG CAATGGGCTT ACCTGGGGGC TCGTTCCTCTG 2760
 TCCCAGAGG AAGCCCCCTC TCCTTCTCCA TGGGCATGAC TCTCCTTCCA GGCCACCACG 2820
 TTTATCTCAC AATGATGTGT TTTGCCTGAC TTTCCCTTTG CGCTGTCTCG TGGGAAAGGT 2880
 CATTCTGTCT GAGACCCAG CTCCTTCTCC AGCTTTGGCT GCGGGCATGG CTTGAGCTTT 2940
 CTGGAGAGCC TCTGCAGGGG GTTGGCCATC AGGGCCCTGT GGCTGGGTCT GCTGCAGAGC 3000
 60 TCCTTGGCTA TCAGGAGAAT CCTGGACACT GTACTGTGCC TCCCAGTTTA CAAACACGCC 3060
 CTTCATCTCA AGTGGCCCTT TAAAGGCCCT GCTGCCATGT GAGAGCTGTG AACAGCTCAG 3120
 CTCTGAGTCG GCAGACTGGG GCTTCTCTCT GGGCCACCAG ATGGAAGGGG GGTATTGTTT 3180
 GCCTCACTCC TGGATGTCTG GTTTAAAGGA AGTGAGTGAG AAAGAATGTG CCAAGATACC 3240
 TGGCTCCTGT GAAACCAGCC TCAGGAGGGA AACTGGGAGA GAGAAGCTGT GGTCTCCTGC 3300
 65 TACATGCCCT GGGAGCTGGA AGAGAAAAAC ACTCCCTAA ACAATCGCAA AATGATGAAC 3360
 CATCATGGGC CACTGTCTC TTTGAGGGGA CAGGTTTAGG GGTTGCGTT CGCCCTTGTG 3420
 GGCTGAAGCA CTAGCTTTT GGTAGCTAGA CACATCTGAC ACCCAAAGGT TCTCTACAAA 3480
 GGCCCAAGAT TGTTTGTAAA GCACTTTGAC TCTTACCTGG AGGCCGCTC TCTAAGGGCT 3540
 TCCTGCGCTC CCACCTCATC TGTCCCTGAG ATGCAGAGCA GGATGGAGGG TCTGCTCTA 3600
 70 GCTCAGCTGT TTCTCCTTGA GGTGCGGAG GAATTGAATT GAATGGGACA GAGGCGAGGT 3660
 GCTGTGGCCA AGAAGATCTC CGAGCAGCAG TGACGGGGCA CTTGTCTGTG TGCTCTCTGG 3720
 GCATGTTAAC CCTTCTGTGG GGCCAAAGGT TTGCATCGTG GATCCAGCTG TGCTCCAGTC 3780
 TGTCCCTCTC TGACTCCACT GACTGCCAC GCCCGGACC AGCAGCTTGG GGACCCTCCA 3840
 GGGTACTAAT GGGGCTCTGT TCTGAGATGG ACAAATTCAG TGTGGAAAT ACATGTTGTA 3900
 75 CTATGCACCT CCCATCTCC TAGGGTTAGG AATAGTTTCA AACATGATTG GCAGACATAA 3960
 CAACGGCAAA TACTCGGACT GGGGCATAGG ACTCCAGAGT AGGAAAAAGA CAAAAGATTT 4020
 GGCAGCTCTA CACAGGCAAC CTACCCCTCT CTCTCCAGCC TCTTATGAA ACTGTTTGT 4080
 TGCCAGTCTT GCCTAAGGC AGAAGATGAA TTGAAGATGC TGTGCATGTT TCCTAAGTCC 4140
 TTGAGCAATC ATGGTGGTGA CAATTGCCAC AAGGGATATG AGGCCAGTGC CACCAGAGGG 4200

TGGTGCCAAG TGCCACATCC CTTCCGATCC ATTCCCCTCT GTATCCTCGG AGCACCCCAG 4260
 TTTGCCCTTG ATGTGTCCGC TGTGTATGTT AGCTGAACTT TGATGAGCAA AATTTCCTGA 4320
 GCGAAACACT CCAAAGAGAT AGGAAAACTT GCCGCCTCTT CTTTITTTGTC CCTTAATCAA 4380
 ACTCAAATAA GCTTAAAAAA AATCCATGGA AGATCATGGA CATGTGAAAT GAGCATTTTT 4440
 TTTCTTTCTT TTTTTTTTTT TTTTITTAAC AAAGTCTGAA CTGAACAGAA CAAGACTTTT 4500
 TCCTCATACA TCTCCAAATT GTTTAACTT ACTTTATGAG TGTITGTTA GAAGTTCGGA 4560
 CCAACAGAAA AATGCAGTCA GATGTCATCT TGGGAATTGGT TICTAAAAAGA GTAAGGCATG 4620
 TCCCTGCCCA GAAACTTAGG AAGCATGAAA TAAATCAAAT GTTTATTTTC CTTCTATT 4680
 AAAATCATGC TAATGCAACA GAAATAGAGG GTTTGTGCCA AATGCTATGA ACGGCCCTTT 4740
 CTAAAGACA AGCAAGGGAG ATTGATATAT GTACAATTG CTCTCATGTT TTT

SEQ ID NO:2 BCU4 Protein sequence:
 Protein Accession #: NP_079191.1

1 11 21 31 41 51
 MSQESDNNKR LVALVPMPSD PPFNTRRAYT SEDEAWKSYL ENPLTAATKA MMIINGDEDS 60
 AAALGLLYDY YKVPDRKRL SVSKASDSQE DQEKRNCLGT SEAQSNLSGG ENRVQVLKTV 120
 PVNLSLNQDH LENSREQYS ISFESSAH PVSIGITVKA EDFTPVFMAP PVHYPRGDGE 180
 EQRVVFIEQT QYDVPSLATH SAYLKDDQRS TPDSTYSESF KDAATEKFRS ASVGAEEMY 240
 DQTSSTGTFQY TLEATKSLRQ KQEGPMTYL NKGQFYAITL SETGDNKCFR HPISKVRSVV 300
 MIVFSEDKNR DEQLKYWKYW HSRQHTAKQR VLDIADYKES FNTIGNIEEI AYNAVSTWD 360
 VNEEAKIFIT VNCLSTDFSS QKGVKGLPLM IQIDTYSYNN RSNKPIHRAV CQIKVFCDKG 420
 AERKIRDEEQ QKRNKNGKGQ ASQTCNNSD DGKLAAPLQ KKSIDITYFKT MPDLHSQPVL 480
 FIPDVHFANL QRTGQVYYNT DDEREGSVL VKRMFRPMEE EFGPVPSKQM KEEGTRVLL 540
 YVRKETDDVF DALMLKSFTV MGLMEAISEK YGLPVEKIAK LYKSKSGKIL VNMDNIEH 600
 YSNEDTIFLN MESMVEGFKV TLMEI

SEQ ID NO:3 BCU7 DNA SEQUENCE VARIANT 1:

Nucleic Acid Accession #: AA428062
 Coding sequence: 1-777 (entire sequence represents open reading frame)

1 11 21 31 41 51
 ATGATAGCAA TCTCTGCCGT CAGCAGTGCA CTCTGTCTCT CCCTTCTCTG TGAAGCAAGT 60
 ACCGTCGTCC TACTCAATTC CACTGACTCA TCCCCGCCAA CCAATAATTT CACTGATATT 120
 GAAGCAGCTC TGAAGCACA ATTAGATTCA GCGGATATCC CCAAAGCCAG GCGGAAGCGC 180
 TACATTTTCG AGAATGACAT GATCGCCATT CTGATTATC ATAATCAAGT TCGGGGCAAA 240
 GTGTTCCAC CGGCAGCAAA TATGGAATAT ATGGTTTGGG ATGAAAATCT TGCAAAATCG 300
 GCAGAGGCTT GGGCGGCTAC TTGCATTGG GACCATGGAC CTCTTACTT ACTGAGATT 360
 TTGGGCCAAA ATCTATCTGT ACGCACTGGA AGATATCGCT CTATTCTCCA GTTGGTCAAG 420
 CCATGGTATG ATGAAGTGA AGATTATGCT TTTCCATATC CCCAGGATTG CAACCCAGAA 480
 TGTCTATGA GATGTTTGG TCCCATGTGC ACACATTATA CGCAGATGGT TTGGGCCACT 540
 TCCAATCGGA TAGGATGCGC AATTCATGCT TGCCAAAACA TGAATGTTG GGGATCTGTG 600
 TGGCGACGTG CAGTTTACTT GGTATGCAAC TATGCCCAA AGGGCAATTG GATTGGAGAA 660
 GCACCATATA AAGTAGGGGT ACCATGTTCA TCTTGTCCTC CAAGTTATGG GGGATCTTGT 720
 ACTGACAATC TGTGTTTTC AGGAGTTACG TCAAACTACC TGTACTGGTT TAAATAA

SEQ ID NO:4 BCU7 DNA SEQUENCE VARIANT 2:

Nucleic Acid Accession #: AA428062
 Coding sequence: 1-777 (entire sequence represents open reading frame)

1 11 21 31 41 51
 ATGATAGCAA TCTCTGCCGT CAGCAGTGCA CTCTGTCTCT CCCTTCTCTG TGAAGCAAGT 60
 ACCGTCGTCC TACTCAATTC CACTGACTCA TCCCCGCCAA CCAATAATTT CACTGATATT 120
 GAAGCAGCTC TGAAGCACA ATTAGATTCA GCGGATATCC CCAAAGCCAG GCGGAAGCGC 180
 TACATTTTCG AGAATGACAT GATCGCCATT CTGATTATC ATAATCAAGT TCGGGGCAAA 240
 GTGTTCCAC CGGCAGCAAA TATGGAATAT ATGGTTTGGG ATGAAAATCT TGCAAAATCG 300
 GCAGAGGCTT GGGCGGCTAC TTGCATTGG GACCATGGAC CTCTTACTT ACTGAGATT 360
 TTGGGCCAAA ATCTATCTGT ACGCACTGGA AGATATCGCT CTATTCTCCA GTTGGTCAAG 420
 CCATGGTATG ATGAAGTGA AGATTATGCT TTTCCATATC CCCAGGATTG CAACCCAGAA 480
 TGTCTATGA GATGTTTGG TCCCATGTGC ACACATTATA CGCAGATGGT TTGGGCCACT 540
 TCCAATCGGA TAGGATGCGC AATTCATGCT TGCCAAAACA TGAATGTTG GGGATCTGTG 600
 TGGCGACGTG CAGTTTACTT GGTATGCAAC TATGCCCAA AGGGCAATTG GATTGGAGAA 660
 GCACCATATA AAGTAGGGGT ACCATGTTCA TCTTGTCCTC CAAGTTATGG GGGATCTTGT 720
 ACTGACAATC TGTGTTTTC AGGAGTTACG TCAAACTACC TGTACTGGTT TAAATAA

SEQ ID NO:5 BCU7 Protein sequence Variant 1:
 Protein Accession #: none

1 11 21 31 41 51
 MIAISAVSSA LLFSLLEAS TVVLLNSTDS SPPTNNFTDI EAALKAQLDS ADIPKARRKR 60

YISQNDMIAI LDYHNQVRGK VFPPAANMEY MVWDENLAKS AEAWAATCIW DHGPSYLLRF 120
 LGQNLISVRTG RYRSILQLVK FWYDEVKDYA FPYPQDCNPR CPMRCFGPMC THYTMVWAT 180
 SNRIGCAIHA CQNMNVWGSV WRRAVYLVCN YAPKGNWIGE APYKVGVPSC SCPPSYGGSC 240
 TDNLCPFPGVT SNLYLYWFK

SEQ ID NO:6 BCU7 Protein sequence Variant 2:
 Protein Accession #: none

1 11 21 31 41 51
 | | | | |
 MIAISAVSSA LLFSLLCEAS TVVLLNSTDS SPPTNNFTDI EAALKAQLDS ADIPKARRKR 60
 YISQNDMIAI LDYHNQVRGK VFPPAANMEY MVWDENLAKS AEAWAATCIW DHGPSYLLRF 120
 LGQNLISVRTG RYRSILQLVK FWYDEVKDYA FPYPQDCNPR CPMRCFGPMC THYTMVWAT 180
 SNRIGCAIHT CQNMNVWGSV WRRAVYLVCN YAPKGNWIGE APYKVGVPSC SCPPSYGGSC 240
 TDNLCPFPGVT SNLYLYWFK

SEQ ID NO:7 BCX2 DNA SEQUENCE

Nucleic Acid Accession #: NM_003014
 Coding sequence: 238-1278 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 GGCGGGITCG CGCCCGAAG GCTGAGAGCT GCGCTGCTC GTGCCCTGTG TGCCAGACGG 60
 CGGAGCTCCG CGGCCGAC CCGCGGCCCC GCTTTGCTGC CGACTGAGT TTGGGGGAAG 120
 AAATCTCTCT GCGCCCGAGA AGATTCTCTC CTCGGCGAAG GGACAGCGAA AGATGAGGGT 180
 GGCAGGAAGA GAAGGCGCTT TCTGTCTGCC GGGGTGCGAG CGCGAGAGGG CAGTGCCATG 240
 TTCTCTCTCA TCTAGTGGC GCTGTGCTG TGGCTGCACC TGGCGCTGGG CGTGC GCGGC 300
 GCGCCCTGCG AGGCGGTGCG CATCCCTATG TGCCGGCACA TGCCCTGGAA CATCACGCGG 360
 ATGCCCAACC ACCTGCACCA CAGCAGCAG GAGAACGCCA TCCTGGCCAT CGAGCAGTAC 420
 GAGGAGCTGG TGGACGTGAA CTGCAGCGCC GTGCTGCGCT TCTTCTCTG TGCCATGTAC 480
 GCGCCCATTT GCACCCATCAT GTTCTGTCAC GACCCTATCA AGCCGTGCAA GTCGGTGTGC 540
 CAACGCGCGC GCGACGACTG CGAGCCCTC ATGAAGATGT ACAACCACAG CTGGCCCGAA 600
 AGCCTGGCCT GCGACGAGCT GCCTGTCTAT GACCGTGGCG TGTGCATTTC GCCTGAAGCC 660
 ATCGTCACGG ACCTCCCGGA GGATGTTAAG TGGATAGACA TCACACCAGA CATGATGGTA 720
 CAGGAAAGGC CTCTTGATGT TGAATGTAAA CGCCTAAGCC CCGATCGGTG CAAGTGATAA 780
 AAGGTGAAGC CACTTTTGGC AACGTATCTC AGCAAAAACT ACAGCTATGT TATTCATGCC 840
 AAAATAAAAG CTGTGCAGAG GAGTGGCTGC AATGAGGTCA CAACGGTGGT GGATGTAAAA 900
 GAGATCTTCA AGTCTCATC ACCATCCCT CGAATCAAG TCCCGCTCAT TACAAATTCT 960
 TCTTGCCAGT TGCCACATC CCTGCCCAT CAAGATGTTT TCATCATGTG TTACGAGTGG 1020
 CGTTCAAGGA TGATGCTTCT TGAATAATGC TTAGTTGAAA AATGGAGAGA TCAGCTTAGT 1080
 AAAAGATCCA TACAGTGGGA AGAGAGGCTG CAGGAACAGC GGAGAACAGT TCAGGACAAG 1140
 AAGAAACAG CCGGGCGCAC CAGTCGTAGT AATCCCCCA AACCAAGGG AAAGCCTCCT 1200
 GCTCCCAAC CAGCCAGTCC CAAGAAGAAC ATTAATACTA GGAGTGCCCA GAAGAGAACA 1260
 AACCAGAAAA GAGTGTGAGC TAAGTAGTTT CCAAAGCGGA GACTTCCGAC TTCCTTACAG 1320
 GATGAGGCTG GGCATTGCCT GGCACAGCCT ATGTAAGGCC ATGTGCCCTT TGCCCTAACA 1380
 ACTCACTGCA GTGCTCTTCA TAGACATATC TTGCAGCATT TTCTTAAAG CTATGCTTCA 1440
 GTTTTCTTT GTAAGCCATC ACAAGCCATA GTGGTAGGTT TGCCCTTGG TACAGAAGGT 1500
 GAGTTAAAGC TGGTGAAAAA GGCTTATTGC ATTGCATTCA GAGTAACCTG TGTGCATACT 1560
 CTAGAAGAGT AGGGAATAA ATGCTTGTTA CAATTCGACC TAATATGTGC ATTGTAAAAAT 1620
 AAATGCCATA TTCAAACAA AACACGTAAT TTTTITACAG TATGTTTTAT TACCTTTTGA 1680
 TATCTGTTGT TGCAATGTTA GTGATGTTT AAAATGTGAT GAAAATATAA TGTTTTTAAG 1740
 AAGGAACAGT AGTGGAAATG ATGTTAAAAA ATCTTTATGT GTTTATGGTC TGCAGAAGGA 1800
 TTTTGTGAT GAAAGGGGAT TTTTGAATAA ATTAGAGAAG TAGCATATGG AAAATTATAA 1860
 TGTGTTTTT TACCAATGAC TTCAGTTTCT GTTTTATGCT AGAACTTAA AAACAAAAAT 1920
 AATAATAAAG AAAATAAAT AAAAAGGAGA GGCAGACAAT GTCTGGATTG CTGTTTTTTG 1980
 GTTACCTGAT TTCCATGATC ATGATGCTTC TTGTCAACAC CCTCTTAAGC AGCACCAGAA 2040
 ACAGTGAGTT TGCTGTACC ATTAGGAGTT AGGTACTAAT TAGTTGGCTA ATGCTCAAGT 2100
 ATTTTATACC CACAAGAGAG GTATGTCACT CATCTTACTT CCCAGGACAT CCACCTGAG 2160
 AATAATTGA CAAGCTTAAA AATGGCCTTC ATGTGAGTGC CAAATTTTGT TTTCTTCAT 2220
 TTAAATATTT TCTTTGCCA AATACATGTG AGAGGAGTTA AATATAAATG TACAGAGAGG 2280
 AAAGTTGAGT TCCACCTCTG AAATGAGAAT TACTTGACAG TTGGGATACT TTAATCAGAA 2340
 AAAAAGAACT TATTGACAG ATTTTATCAA CAAATTTTCA ATTTGTGGAC AATTGGAGGC 2400
 ATTTATTTTA AAAACAAT TTAATGGCCT TTTGCTAACA CAGTAAGCAT GTATTTTATA 2460
 AGGCATTCAA TAAATGCACA ACGCCCAAAG GAAATAAAAA CCTATCTAAT CCTACTCTCC 2520
 ACTACACAGA GGTAATCACT ATTAGTATTT TGGCATATTA TTCTCCAGGT GTTTGCTTAT 2580
 GCACITATAA AATGATTTGA ACAAATAAAA CTAGGAACCT GTATACATGT GTTTCATAAC 2640
 CTGCCTCTT TGCTTGGCCC TTTATTGAGA TAAGTTTCC TGCAAGAAA GCAGAAACCA 2700
 TCTCATTTCT AACAGCTGTG TTATATTCCA TAGTATGCAT TACTCAACAA ACTGTTGTGC 2760
 TATTGGATAC TTAGGTGGT TCTCACTGA CAATACTGAA TAAACATCTC ACCGGAATTC

SEQ ID NO:8 BCX2 Protein sequence:
 Protein Accession #: NP_003005.1

1 11 21 31 41 51
 | | | | |
 MFLSILVALC LWLHLALGVR GAPCEAVRIP MCRHMPWNIT RMPNHLHHST QENAILAIEQ 60

YEELVDVNCs AVLRFFFCAM YAPICILEFL HDPIKPKSV QQRARDCEP LMKMYNHSWP 120
 ESLACDELVP YDRGVCISPE AIVTDLPELV KWIDITPDM VQERPLDVDC KRLSPDRCKC 180
 KKVKPTLATY LSKNYSYVIH AKIKAVQRSG CNEVTTVVDV KEIFKSSSPI PRTQVPLITN 240
 SSCQCPHILP HQDVLIMCYE WRSRMMLEN CLVEKWRDQL SKRSIQWEER LQEQRRTVQD 300
 KKKTAGRTSR SNPPKPKGKP PAPKPASPKK NIKTRSAQKR TNPKRV

SEQ ID NO:9 CBK1 DNA SEQUENCE

Nucleic Acid Accession #: NM_032391

Coding sequence: 129-302 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 15 GTCTTCTCTC TCCTAGCCTA AGGCGTGCAA ACAGAGCGCC ACTGGGAGGC TGAAACCTTT 60
 AGGCCGATGC TTGCTTGCAA GGTCAAGCAA GCTGGATCTT GGTCCCCACC TTTCAGAGA 120
 GAACAGCGAT GTTGTGCGCC CATTTCTCAG ATCAAGGACC GGCCCATCTT ACTACCTCCA 180
 AGAGTGCCTT TCTCTCTAAT AAGAAAACAT CTACTTTGAA ACATCTACTG GCGAGACCA 240
 20 GGAGTGATGG CTCAGCCTGT AATTCTGGAA TTTCGGGAGG CCGAGGCAGG AAGATTCTTT 300
 GAGCACAGGA GTTCCAGACC AGCCTGGGCA ATGTAGCAAG ACGCTGTCTC TATTTATACA 360
 ATAAATTTT TTTAAAAAG G

SEQ ID NO:10 CBK1 Protein sequence:

Protein Accession #: NP_115767

1 11 21 31 41 51
 | | | | |
 30 MLCAHFSDQG PAHLTTSKSA FLSNKKSTL KHLGETRSD GSACNSGISG GRGRKIP

SEQ ID NO:11 CHA1 DNA SEQUENCE

Nucleic Acid Accession #: NM_020182

Coding sequence: 96-854 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 40 TCCTTGGGTT CGGGTGAAGG CGCCTGGGGG TCGTGGCCA TGATCCCCGA GCTGCTGGAG 60
 AACTGAAGGC GGACAGTCTC CTGCGAAACC AGGCAATGGC GGAGCTGGAG TTTGTTTCAGA 120
 TCATCATCAT CGTGGTGGTG ATGATGGTGA TGGTGGTGGT GATCACGTGC CTGCTGAGCC 180
 ACTACAAGCT GTCTGCACGG TCCTTCATCA GCCGGCACAG CCAGGGGGCG AGGAGAGAAG 240
 45 ATGCCCTGTC CTCAGAAGGA TGCCCTGTGGC CCTCGGAGAG CACAGTGTCA GGCAACGGAA 300
 TCCAGAGGCC GCAGGTCTAC GCCCCGCCCTC GGCCCAACGA CCGCCTGGCC GTGCCGCCCT 360
 TCGCCCAAGC GGAGCGCTTC CACCGCTTCC AGCCACCTA TCCGTACCTG CAGCACGAGA 420
 TCGACCTGCC ACCCACCATC TCGCTGTGAG ACGGGGAGGA GCCCCACACC TACCAGGGCC 480
 CCTGCACCTT CCAGCTTCGG GACCCCGAGC AGCAGCTGGA ACTGAACCGG GAGTCGGTGC 540
 50 GCGCACCCCC AAACAGAAACC ATCTTCGACA GTGACCTGAT GGATAGTGCC AGGCTGGGCG 600
 GCCCCTGCCC CCCAGCAGT AACTCGGGCA TCAGCGCCAC GTGCTACGGC AGCGGCGGGC 660
 GCATGGAGGG GCCGCCGCCC ACCTACAGCG AGGTCAATCG CCACTACCCG GGGTCTCTCT 720
 TCCAGACCCA GCAGAGCAGT GGGCCGCCCT CCTTGCTGGA GGGGACCCGG CTCACACACA 780
 CACACATCGC GCCCCTAGAG AGGCGAGCCA TCTGGAGCAA AGAGAAGGAT AAACAGAAAG 840
 55 GACACCTCTT CTAGGGTCCC CAGGGGGGCC GGGCTGGGGC TCGGTAGGTG AAAAGGCAGA 900
 AACTCCGCG CTCTCTAGAA GAGGAGTGAG AGGAAGGCGG GGGCGCGAGC AACGCATCGT 960
 GTGGCCCTCC CCTCCACCT CCTGTGTAT AAATATTTAC ATGTGATGTC TGGTCTGAAT 1020
 GCACAAGCTA AGAGAGCTTG CAAAAAAGAA AAGAAAAAAG AAAAAAAGAA ACCACGTTTC 1080
 TTTGTTGAGC TGTGCTCTGA AGGCAAAAGA AAAAAAATTT CTACAGTAAA AAAAAAAGAA 1140
 A

SEQ ID NO:12 CHA1 Protein sequence:

Protein Accession #: NP_064567

1 11 21 31 41 51
 | | | | |
 65 MALEFVQII IIVVMMVMV VVITCLLSHY KLSARFISR HSQGRRREDA LSSEGCLWPS 60
 ESTVSGNGIP EPQVYAPRP TDRLAVPPFA QRERFHRFQ TYPYLQHEID LPPTISLSDG 120
 70 EEPFPYQGPC TLQLRDEEQ LELNRESVRA PPNRTIFDSD LMSARLGGP CPPSSNSGIS 180
 ATCYGSGGRM EGPPTYSEV IGHYPGSSFQ HQQSSGPPSL LEGTRLHHTH IAPLESAAIW 240
 SKEKDKQKGH PL

SEQ ID NO:13 CJA5 DNA SEQUENCE

Nucleic Acid Accession #: NM_012445

Coding sequence: 276-1271 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |

5
10
15
20
25
30

```

GCACGAGGGA AGAGGGTGAT CCGACCCGGG GAAGGTCGCT GGGCAGGGCG AGTTGGGAAA 60
GCGGCAGCCC CCGCCGCCCC CGCAGCCCCT TCTCCTCCTT TCTCCCACGT CCTATCTGCC 120
TCTCGCTGGA GGCAGGCGCG TGCAGCATCG AAGACAGGAG GAACCTGGAGC CTCATTGGCC 180
GGCCCGGGGG GCGCGGCTCG GGCTTAAATA GGAGCTCCGG GCTCTGGCTG GGACCCGACC 240
GGTGCAGGCC GCGCTCCCGC TGCTCTCTGC GGGTGATGGA AAACCCACAG CCGGCGCGCG 300
CCCTGGGCAA GGCCCTCTGC GCTCTCTCTC TGGCCACTCT CGGCGCCGCC GGCCAGCCTC 360
TTGGGGGAGA GTCCATCTGT TCCGCCAGAG CCCCAGCCAA ATACAGCATC ACCTTCACGG 420
GCAAGTGGAG CCAGACGGCC TTCCCCAAGC AGTACCCCTT GTTCCGCCCC CCTGCGCAGT 480
GGTCTTCGCT GCTGGGGGCC GCGCATAGCT CCGACTACAG CATGTGGAGG AAGAACCAGT 540
ACGTCAGTAA CGGGCTGCGC GACTTTGCGG AGCGCGGCGA GGCTTGGGCG CTGATGAAGG 600
AGATCGAGGC GCGGGGGGAG GCGCTGCAGA GCGTGCACGC GGTGTTTTCG GCGCCCGCCG 660
TCCCCAGCGG CACCGGGCAG ACGTCGGCGG AGCTGGAGGT GCAGCGCAGG CACTCGCTGG 720
TCTCGTTTGT GGTGCGCATC GTGCCAGGCC CCGACTGGTT CGTGGGCGTG GACAGCCTGG 780
ACCTGTGCGA CGGGGACCGT TGGCGGGAAC AGGCGGCGCT GGACCTGTAC CCCCTACGACG 840
CCGGGACGGA CAGCGGCTTC ACCTTCTCCT CCCCACACTT CGCCACCATC CCGCAGGACA 900
CGGTGACCGA GATAACGTCC TCCTCTCCCA GCCACCCGGC CAACCTCTTC TACTACCCGC 960
GGCTGAAGGC CCTGCTCTCC ATCGCCAGGG TGACACTGGT GCGGCTGCGA CAGAGCCCCA 1020
GGGCTTTCAT CCTTCCCGCC CAGTCTCTGC CCAGCAGGGA CAATGAGATT GTAGACACGG 1080
CCTCAGTTCC AGAAACGCGC CTGGACTGCG AGGTCTCCCT GTGGTCGTCC TGGGGACTGT 1140
CGGAGGCGCA CTGTGGGAGG CTCGGGACCA AGAGCAGGAC TCGCTACGTC CCGGTCCAGC 1200
CCGCCAACAA CGGGAGCCCC TGCCCGGAGC TCGAAGAAGA GGCTGAGTGC GTCCCTGATA 1260
ACTGCGTCTA AGACCAGAGC CCGCAGGCC CTGGGGCCCC CGGAGCCATG GGGTGTGCGG 1320
GGCTCTCTGT CAGGCTCATG CTGCAGGCGG CCGAGGCACA GGGGGTTTCG CGCTGCTCCT 1380
GACCGCGGTG AGGCGCGGCC GACCATCTCT GCACTGAAGG GCCCTCTGGT GCGCGGCACG 1440
GGCATTGGGA AACAGCTTCC TCCTTTCCCA ACCTTGCTTC TTAGGGGCCC CCGTGTCCCG 1500
TCTGCTCTCA GCCTCTCTCT CTGTCAGGAT AAAGTCATCC CCAAGGCTCC AGCTACTCTA 1560
AATTATGGTC TCCTTATAAG TTATTGCTGC TCCAGGAGAT TGTCTTTCAT CGTCCAGGGG 1620
CCTGGTCTCC ACGTGGTTGC AGATACCTCA GACCTGGTGC TCTAGGCTGT GCTGAGCCCA 1680
CTCTCCCGAG GCGCGATCCA AGCGGGGGCC ACTTGAGAAG TGAATAAATG GGGCGGTTTC 1740
GGAAGCGTCA GTGTTTCCAT GTTATGGATC TCTCTGCGTT TGAATAAAGA CTATCTCTGT 1800
TGCTCAC

```

35
SEQ ID NO:14 CJA5 Protein sequence:
Protein Accession #: NP_036577

40
45

```

1      11      21      31      41      51
|      |      |      |      |      |
MENPSPAAL GKALCALLL TLGAAGQPLG GESIC SARAP AKYSITFTGK WSQTAFPKQY 60
PLFRPPAQWS SLLGAAHSSD YSMWRKNQYV SNGLRDFAE GEAWALMKEI EAAAGEALQSV 120
HAVFSAPAVP SGTGQTSACL EVQRRHSLVS FVVRIVPSPD WFGVDSLSDL CDGDRWREQA 180
ALDLYPYDAG TDSGTFSSP NFATIPQDIV TEITSSSPSH PANSFYYPRL KALPPIARVT 240
LVRLRQSPRA FIPPAVLPS RDNEIVDSAS VPETPLDCEV SLWSSWGLCG GHCGRLLGTS 300
RTRYVRVQPA NNGSPCELE EEAECVPDNC V

```

50
SEQ ID NO:15 LBH9 DNA SEQUENCE
Nucleic Acid Accession #: NM_002391
Coding sequence: 26-457 (underlined sequences correspond to start and stop codons)

55
60
65

```

1      11      21      31      41      51
|      |      |      |      |      |
CGGGCGAAGC AGCGCGGGCA GCGAGATGCA GCACCGAGGC TTCTCTCTCC TCACCTCTCT 60
CGCCCTGCTG GCGCTCACCT CCGCGGTGCG CAAAAAGAAA GATAAGGTGA AGAAGGGCGG 120
CCCGGGGAGC GAGTGCCTGT AGTGGGCTCG GGGGCCCTGC ACCCCAGCA GCAAGGATTG 180
CGGCGTGGGT TTCCGCGAGG GCACCTGCGG GGCCAGACC CAGCGCATCC GGTGCAGGGT 240
GCCCTGCAAC TGGAAGAAGG AGTTTGAGAG CGACTGCAAG TACAAGTTTG AGAACTGGGG 300
TGCGTGTGAT GGGGGCACAG GCACCAAGT CCGCCAAGGC ACCCTGAAGA AGGCGCGCTA 360
CAATGCTCAG TGCCAGGAGA CCATCCGCGT CACCAAGCCC TGCACCCCA AGACCAAGC 420
AAAGGCCAAA GCCAAGAAAG GGAAGGGAAA GGACTAGACG CCAAGCCTGG ATGCCAAGGA 480
GCCCTGGTG TCACATGGGG CTTGGCCACG CCTTCCCTCT CCCAGGCCCG AGATGTGACC 540
CACCAGTGCC TTCTGTCTGC TCGTTAGCTT TAATCAATCA TGCCCTGCCT TGTCCCTCTC 600
ACTCCCAGC CCCACCCCTA AGTGCCCAAA GTGGGGAGGG ACAAGGGATT CTGGGAAGCT 660
TGAGCCTCCC CCAAAGCAAT GTGAGTCCCA GAGCCCGCTT TTGTCTCTCC CCACAATTCC 720
ATTACTAAGA AACACATCAA ATAACTGAC TTTTCCCCC CAATAAAAGC TCTTCTTTT 780
TAATAT

```

70
SEQ ID NO:16 LBH9 Protein sequence:
Protein Accession #: NP_002382

75

```

1      11      21      31      41      51
|      |      |      |      |      |
MQHRGFLLLT LLALLAL TSA VAKKDKVKK GPGSECAEW AWGPCTPSSK DCGVFREGT 60
CGAQTQIRIC RVPCNWKKEF GADCKYKFEN WGACDGGTGT KVRQGTLLKA RYNAQCQETI 120
RVTKPTPKT KAKAKAKKGK GKD

```

SEQ ID NO:17 LEM9 DNA SEQUENCE

Nucleic Acid Accession #: NM_005244

Coding sequence: 1-1617 (underlined sequences correspond to start and stop codons)

5
10
15
20
25
30
35

```

1      11      21      31      41      51
|      |      |      |      |      |
ATGGTAGAAC TAGTGATCTC ACCCAGCCTC ACTGTAAACA GCGATTGTCT GGATAAACTG 60
AAGTTTAACC GTGCTGACGC TGCTGTGTGG ACTCTGAGTG ACAGACAAGG CATCACCAAA 120
TCGGCCCCCC TGAGAGTGTC CCAGCTCTTC TCCAGATCTT GCCCACGTGT CCTCCCCCGC 180
CAGCCTTCCA CAGCCATGGC AGCCTACGGC CAGACGCAGT ACAGTGCGGG GATCCAGCAG 240
GCTACCCCTT ATACAGCTTA CCCACCTCCA GCACAAGCCT ATGGAATCCC TTCTACAGC 300
ATCAAGACAG AAGACAGCTT GAACCATTCC CCTGGCCAGA GTGGATTCTT CAGCTATGGC 360
TCCAGCTTCA GCACCTCACC CACTGGACAG AGCCCATACA CCTACCAGAT GCACGGCACA 420
ACAGGTTTCT ATCAAGGAGG AAATGGACTG GGCAACGCAG CCGGTTTCGG GAGTGTGCAC 480
CAGGACTATC CTTCCTAGCT CGGCTTCCCC CAGAGCCAGT ACCCCAGTA TTACGGCTCA 540
TCCTACAACC CTCCCTACGT CCCGGCCAGC AGCATCTGCC CTTCGCCCTT CTCCACGTCC 600
ACCTACGTCC TCCAGGAGGC ATCTCACAAAC GTCCCAACC AGAGTTCCGA GTCACCTTGT 660
GGTGAATACA ACACACACAA TGGACCTTCC ACACCAGCGA AAGAGGGAGA CACAGACAGG 720
CCGCACCGGG CCTCCGACGG GAAGCTCCGA GGCCGGTCTA AGAGGAGCAG TGACCCGTCC 780
CCGGCAGGGG ACAATGAGAT TGAGCGTGTG TTCGTGTGGG ACTTGGATGA GACAATAATT 840
ATTTTTCACT CCTTACTCAC GGGGACATTT GCATCCAGAT ACGGGAAGGA CACCACGACG 900
TCCGTGCGCA TTGGCCTTAT GATGGAAGAG ATGATCTTCA ACCTTGCGA TACACATCTG 960
TTCTTCAATG ACCTGGAGGA TTGTGACCAG ATCCACGTGT ATGACGTCTC ATCAGATGAC 1020
AATGGCCAAG ATTTAAGCAC ATACAACCTT TCCGCTGACG GCTTCCACAG TTCGGCCCA 1080
GGAGCCAACC TGTGCCTGGG CTCTGGCGTG CACGGCGCGG TGGACTGGAT GAGGAAGCTG 1140
GCCTTCCGCT ACCGGCGGGT GAAGGAGATG TACAATACCT ACAAGAACAA CGTTGGTGGG 1200
TTGATAGGCA CTCCCAAAG GGAGACCTGG CTACAGCTCC GAGCTGAGCT GGAAGCTCTC 1260
ACAGACCTCT GGCTGACCCA CTCCCTGAAG GCACTAAACC TCATCAACTC CCGCCCAAC 1320
TGTGTCAATG TGCTGGTACC CACCCTCAA CTAATTCCTG CCTGGCCAA AGTCCTGCTA 1380
TATGGCCTGG GGTCTGTGTT TCCTATTGAG AACATCTACA GTGCAACCAA GACAGGGAAG 1440
GAGAGCTGCT TCGAGAGGAT AATGCAGAGA TTCGGCAGAA AAGCTGTCTA CGTGGTGATC 1500
GGTGATGTG TGGAAGAGGA GCAAGGAGCG AAAAAGCACA ACATGCCTTT CTGGCGGATA 1560
TCCTGCCACG CAGACCTGGA GGCCTGAGG CACGCCCTGG AACTGGAGTA TTTATAG

```

SEQ ID NO:18 LEM9 Protein sequence:

Protein Accession #: NP_005235

40
45
50

```

1      11      21      31      41      51
|      |      |      |      |      |
MVELVISPSL TVNSDCLDKL KFNRAAAVW TLSDRQGITK SAPLRVSQLF SRSCPRLVLR 60
QPSTAMAAYG QTQYSAGTQQ ATPYTAYPPP AQAYGIPSYS IKTEDSLNHS PQQSGFLSYG 120
SSFSTSPGTQ SPYTYQMHTG TGFYQGGNGL GNAAGFGSVH QDYPSYPGFP QSQYPQYYS 180
SYNPPYVPAS SICPSPLSTS TYVLQEAHSH VFNQSSSESLA GEYNTNHGFS TPAKEGDTDR 240
PHRASDGLRL GRSKRSSDES PPDNEIERV FVWDLDETII IFHSLLTGTF ASRYGKDTT 300
SVRIGLMME E MIFNLADTHL PFNDLEDQDQ IHVDDVSSDD NGQDLSTYNF SADGFHSSAP 360
GANLCLGSGV HGGVDWMRKL AFRYRRVKEM YNTYKNNVGG LIGTPKRETW LQLRAELEAL 420
TDLWLTHSLK ALNLDNSRFL CVNVLVTTTQ LIPALAKVLL YGLGSVPFIE NIYSATKTGK 480
ESCFERIMQR FGRKAVYVVI GDGVEEQGA KKHNMFWRI SCHADLEALR HALELEYL

```

SEQ ID NO:19 OAA1 DNA SEQUENCE

Nucleic Acid Accession #: NM_002740

Coding sequence: 178-1968 (underlined sequences correspond to start and stop codons)

55
60
65
70
75
80

```

1      11      21      31      41      51
|      |      |      |      |      |
CCGCGGTTCC GGCTGCTCCG GCGAGGCGAC CCTTGGGTCG GCGCTGCGGG CGAGGTGGGG 60
AGGTAGGTGG GCGGACGGCC GCGGTTCTCC GGCAAGCGCA GCGCGCGGAG TCCCCACGG 120
CGCCCCAAGC GCCCCCCGCA CCCC CGGCTTGA GCGCGGGGAG TGAGGAGATG 180
CCGACCCAGA GGGACAGCAG CACCATGTCC CACACGGTCG CAGCGGCGGG CAGCGGGGAC 240
CATTCCCACC AGGTCCGGGT GAAAGCCTAC TACCGCGGGG ATATCATGAT AACACATTTT 300
GAACTTCCA TCTCCTTTGA GGGCCTTTGC AATGAGGTTT GAGACATGTG TTCTTTTGAC 360
AACGAACAGC TCTTCAACCAT GAAATGGATA GATGAGGAAG GAGACCCGTG TACAGTATCA 420
TCTCAGTTGG AGTTAGAAGA AGCCTTTAGA CTTTATGAGC TAAACAAGGA TTCTGAATC 480
TTGATTCTATG TGTTCCTTTG TGTACCAGAA CGTCTTGGGA TGCCTTGTTC AGGAGAAGAT 540
AAATCCATCT ACCGTAGAGG TGCACGCGCG TGGAGAAAGC TTTATTGTGC CAATGGCCAC 600
ACTTTCCAAG CCAAGCGTTT CAACAGGCGT GCTCACTGTG CCATCTGCAC AGACCGAATA 660
TGGGGACTTG GACGCCAAGG ATATAAGTGC ATCAACTGCA AACTCTTGTT TCATAAGAAG 720
TGCCATAAAC TCGTCACAAT TGAATGTGGG CGGCATCTCT TGCCACAGGA ACCAGTGATG 780
CCCATGGATC AGTCATCCAT GCATCTTGAC CATGCACAGA CAGTAATTCC ATATAATCCT 840
TCAAGTCATG AGAGTTTGGG TCAAGTTGGT GAAGAAAAAG AGGCAATGAA CACCAAGGAA 900
AGTGGCAAAG CTTTATCCAG TCTAGTCTT CAGGATTTTG ATTGTCTCCG GGTAAATAGGA 960
AGAGGAAGTT ATGCCAAAGT ACTGTTGGTT CGATTAAAAA AAACAGATCG TATTTATGCA 1020
ATGAAAGTTG TGAAAAAGA GCTTGTTAAT GATGATGAGG ATATTGATTG GGTACAGACA 1080
GAGAAGCATG TGTGTGAGCA GGCATCCAAT CATCCTTCC TTGTTGGGCT GCATTCTTGC 1140
TTTCAGACAG AAAGCAGATT GTTCTTTGTT ATAGAGTATG TAAATGGAGG AGACCTAATG 1200
TTTCATATGC AGCGACAAAG AAAACTTCTT GAAGAACATG CCAGATTTTA CTCTGCAGAA 1260

```


ATCAGTCTAG CATTAAATTA TCTTCATGAG CGAGGGATAA TTTATAGAGA TTTGAAACTG 1320
 GACAATGTAT TACTGGACTC TGAAGGCCAC ATTAAACTCA CTGACTACGG CATGTGTAAAG 1380
 GAAGGATTAC GGCACAGGAGA TACAACCAGC ACTTTCCTGTG GTACTCCTAA TTACATGTGT 1440
 CCTGAAATTT TAAGAGGAGA AGATTATGGT TTCAGTGTGG ACTGGTGGGC TCTTGGAGTG 1500
 CTCTATGTTTG AGATGATGGC AGGAAGGTCT CCATTTGATA TTGTTGGGAG CTCGGATAAC 1560
 CCTGACCAGA ACACAGAGGA TTATCTCTTC CAAGTTATTT TGGAAAAACA AATTGCGATA 1620
 CCACGTTCTC TGTCTGTAAA AGCTGCAAGT GTTCTGAAGA GTTTTCTTAA TAAGGACCCT 1680
 AAGGAACGAT TGGGTTGTCA TCCTCAAACA GGATTTGCTG ATATTACAGG ACACCCGTTT 1740
 TTCCGAAATG TTGATTGGGA TATGATGGAG CAAAAACAGG TGGTACCTCC CTTTAAACCA 1800
 AATATTCTCT GGGAAATTGG TTTGGACAAC TTTGATCTC AGTTTACTAA TGAACCTGTC 1860
 CAGCTCACTC CAGATGACGA TGACATTGTG AGGAAGATTG ATCAGTCTGA ATTTGAAGGT 1920
 TTTGAGTATA TCAATCCTCT TTTGATGTC GCAGAAGAAT GTGCTGATC CTCATTTTTC 1980
 AACCATGTAT TCTACTCATG TTGCCATTTA ATGCATGGAT AAACCTGCTG CAAGCCTGGA 2040
 TACAATTAAC CATTTTATAT TTGCCACCTA CAAAAAACA CCAATATCT TCTCTGTAG 2100
 ACTATATGAA TCAATTATTA CATCTGTTTT ACTATGAAAA AAAAATTAAT ACTACTAGCT 2160
 TCCAGACAAT CATGTCAAAA TTTAGTTGAA CTGGTTTTTC AGTTTTTAAA AGGCCCTACAG 2220
 ATGAGTAATG AAGTTACCTT TTTTGTTTAA AAAAAAATAA G

SEQ ID NO:20 QAA1 Protein sequence:

Protein Accession #: NP_002731

1 11 21 31 41 51
 MSHTVAGGGS GDHSQVRVK AYYRGDIMIT HFEPSISFEG LCNEVRDMCS FDNEQLFTMK 60
 WIDEEGDPT VSSQLEEEA FRLYELNKDS ELLIHVFPVC PERPGMPCPG EDKSIYRRA 120
 RRWRKLYCAN GHTFQAKRFN RRAHCAICTD RIWGLGRQGY KCINCKLLVH KKCHKLVITIE 180
 CGRHSLPQEP VMPMDQSSMH SDHAQTVIPY NPSSHESLDQ VGEKEAMNT RESGKASSSL 240
 GLQDFDLRV IGRGSYAKVL LVRLKKTDRI YAMKVVKEL VNDEDEDIDW QTEKHVFEQA 300
 SNHPFLVLGH SCFTVESRLF FVIEYVNGGD LMFHMQRQRK LPEEHARFYS AEISLALNLY 360
 HERGIYIDL KLDNVLLDSE GHIKLTDYGM CKEGLRPGDT TSTFCGTFNY IAPFILRGED 420
 YGFSVDWWAL GVLMFEMMAG RSPFDIVGSS DNPQNTEDY LFQVILEKQI RIPRSLSVKA 480
 ASVLKSFNLK DPKERLGCHP QTFADIQGH PFFRNVWDWM MEQKQVPPF KPNISGEFGL 540
 DNFDSQFTNE PVQLTPDDDD IVRKIDQSEF EGFEYINPLL MSAEECV

SEQ ID NO:21 OBH2 DNA SEQUENCE

Nucleic Acid Accession #: L05628

Coding sequence: 197-4792 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 CCAGGCGGCG TTGCGGCCCC GCGCCCGGCT CCTGCGCGCG CCGCGCGCGC CGCCGCGCGC 60
 GCGCGCGCGC CCGCGCGCAG CGCTAGCGCC AGCAGCCGGG CCGGATCACC CGCCGCGCGC 120
 TGCCCGCGCG CGCCCGCGCC AGCAACCGGG CCGGATCACC CGCCGCGCGC TGCCCGCGCG 180
 CGCCCGCGCG ACCGCGCATGG CGCTCCGGGG CTCTGCGAGC GCGGATGGCT CCGACCCGCT 240
 CTGGGACATGG AATGTCACGT GGAATACCAG CAACCCGAC TTCACCAAGT GCTTTTCAAA 300
 CACGGTCCCTC GTGTGGGTGC CTTGTTTTTA CCTCTGGGCC TGTTCCTCTT TCTACTTCTT 360
 CTATCTCTCC CGACATGACC GAGGCTACAT TCAGATGACA CCTCTCAACA AAACCAAAAC 420
 TGCTCTGGGA TTTTTCCTGT GGATCGTCTG CTGGGCGAGC CTCTCTTACT CTTCTTGGGA 480
 AAGAAAGTCGG GGCATATATCC TGGCCCCAGT GTTCTGTGTC AGCCCAACTC TCTTGGGCAT 540
 CACCACGCTG CTTGCTACCT TTTTAATTC A GCTGGAGAGG AGGAAGGGAG TTCAGTCTTC 600
 AGGGATCATG CTACCTTTCT GCGTGGTAGC CCTAGTGTGT GCCCTAGCCA TCCTGAGATC 660
 CAAAATTATG ACAGCCTTAA AAGAGGATGC CCAGGTGGAC CTGTTTCGTG ACATCACTTT 720
 CTACGTCTAC TTTTCCCTCT TACTCATTCA GCTCGTCTTG TCTGTCTTCT CAGATCGCTC 780
 ACCCTGTCTC TCGGAAACCA TCCACGACCC TAATCCCTGC CCAGAGTCCA GCGCTTCTCT 840
 CCTGTGCGAG ATCACCCTCT GGTGGATCAC AGGGTTGATT GTCCGGGGCT ACCGCCAGCC 900
 CCTGGAGGGC AGTGACCTCT GGTCTTAAA CAAGGAGGAC ACGTCGGAAC AAGTCGTGCC 960
 TGTTTTGGTA AAGAACTGGA AGAAGGAATG CGCCAAGACT AGGAAGCAGC CGGTGAAGGT 1020
 TGTGTACTCC TCCAAGGATC CTGCCCAGCC GAAAGAGAGT TCCAAGTGG ATCCGAATGA 1080
 GGAGGTGGAG GCTTTGATCG TCAAGTCCCC ACAGAAGGAG TGAACCCCTT CTCTGTTTAA 1140
 GGTGTATATC AAGACCTTTG GGCCTACTT CCTCATGAGC TTCTTCTTCA AGGCCATCCA 1200
 CGACCTGATG ATGTTTCCG GCGCGCAGAT CTTAAAGTTG CTCATCAAGT TCGTGAATGA 1260
 CACGAAGGCC CCAGACTGGC AGGGCTACTT CTACACCGTG CTGCTGTTTG TCACTGCCTG 1320
 CCTGCAGACC CTCGTGCTGC ACCAGTACTT CCACATCTGC TTCGTCAAGT GCATGAGGAT 1380
 CAAGACCGCT GTCATTGGGG CTGTCTATCG GAAGGCCCTG GTGATCACC AATCAGCCAG 1440
 AAAATCTTCC ACGGTGCGGG AGATTGTCAA CCTCATGTCT GTGGACGCTC AGAGGTTCAT 1500
 GGACTTGGCC ACGTACATTA ACATGATCTG GTCAGCCCC CTGCAAGTCA TCCTTGTCTT 1560
 CTACCTCTCT TGGCTGAATC TGGGCCCTTC CGTCTGTGCT GGAGTGGCGG TGATGGTCTT 1620
 CATGGTGCCC GTCATGCTGT TGATGGCGAT GAAGACCAAG ACGTATCAGG TGCCCCACAT 1680
 GAAGAGCAAA GACATCGGA TCAAGCTGAT GAACGAAATT CTCATGGGA TCAAAAGTGT 1740
 AAAGCTTTAT GCCTGGGAGC TGGCATTCAA GGACAAGGTG CTGGCCATCA GGCAGGAGGA 1800
 GCTGAAGGTG CTGAAGAAGT CTGCCTACCT GTCAGCCGTG GGCACCTTCA CTTGGGTCTG 1860
 CACGCCCTTT CTGGTGGCCT TGTGCACATT TGCCGTCTAC GTGACCATTT ACGAGAACA 1920
 CATCTGTGAT GCCCAGACAG CCTCTGTGTC TTTGGCCTTG TTCAACATCC TCCGCTTTC 1980
 CCTGAACATT TCCCCCATGG TCATCAGCAG CATCGTGCAG GCGAGTGTCT CCCTCAAACG 2040
 CCTGAGGATC TTTCTCTCCC ATGAGGAGCT GGAACCTGAC AGCATCGAGC GACGGCTGT 2100
 CAAAGACGGC GGGGGCAGCA ACAGCATCAC CGTGAGGAAT GCCACATTC CTTGGGCCAG 2160
 GAGCGACCCT CCCACATGTA ATGGCATCAC CTCTCTCATC CCCGAAGGTG CTTTGTGTGG 2220

5 CGTGGTGGGC CAGGTGGGCT GCGGAAAGTC GTCCCTGCTC TCAGCCCTCT TGGCTGAGAT 2280
 GGACAAAGTG GAGGGGACAG TGGCTATCAA GGGCTCCGTG GCCTATGTGC CACAGCAGGC 2340
 CTGGATTGAG AATGATTTCT TCCGAGAAAA CATCCTTTTT GGATGTCAGC TGGAGGAACC 2400
 ATATTACAGG TCCGTGATAC AGGCCTGTGC CCTCCTCCCA GACCTGGAAA TCCTGCCCAG 2460
 TGGGGATCGG ACAGAGATTG GCGAGAAGGG CGTGAACCTG TCTGGGGGCC AGAAGCAGCG 2520
 CGTGAGCCTG GCCCGGGCCG TGTACTCCAA CGCTGACATT TACCTCTTCG ATGATCCCCT 2580
 CTACAGCAGT GATGCCCCAT TGGGAAAACA CATCTTTGAA AATGTGATTG GCCCAAGGG 2640
 GATGCTGAAG AACAAAGACG GATCTTGGT CACGCACAGC ATGAGCTACT TGCCCGCAGT 2700
 10 GGACGTCATC ATCGTCATGA GTGGCGGCAA GATCTCTGAG ATGGGCTCCT ACCAGGAGCT 2760
 GCTGGCTCGA GACGGCCGCT TCGCTGAGTT CCTGCGTACC TATGCCAGCA CAGAGCAGGA 2820
 GCAGGATGCA GAGGAGAACG GGGTCACGGG CGTCAGCGGT CCAGGGAAGG AAGCAAAGCA 2880
 AATGGAGAA TGGCATGCTG TGACGGACAG TGCAGGGAAG CAACTGCAGA GACAGCTCAG 2940
 CAGCTCCTCC TCCTATAGTG GGGACATCAG CAGGCACAC AACAGCACCG CAGAACTGCA 3000
 15 GAAAGCTGAG GCCAAGAAGG AGGAGACCTG GAAGCTGATG GAGGCTGACA AGGCGCAGAC 3060
 AGGGCAGGTG AAGCTTTCCG TGTACTGGGA CTACATGAAG GCCATCGGAC TCTTCATCTC 3120
 CTTCTCAGC ATCTTCTCTT TCAATGTGTA CCATGTGTCC CGCTGGCTT CCAACTATTG 3180
 GCTCAGCCTC TGGACTGATG ACCCATCTGT CAACGGGACT CAGGAGCACA CGAAAGTCCG 3240
 GCTGAGCGTC TATGGAGCCC TGGGCATTTT ACAAGGGATC GCCGTGTTTG GCTACTCCAT 3300
 20 GGCCTGTCTC ATCGGGGGGA TCCTGGCTTC CCGCTGTCTG CACGTGGACC TGTGACACAG 3360
 CATCTCGGG TCACCCATGA GCTCTTTTGA GCGGACCCCG AGTGGGAACC TGGTGAACCG 3420
 CTCTCCCAAG CAGGTGACCA CAGTGGACTC CATGATCCCG GAGGTCATCA AGATGTTTCA 3480
 GGGCTCCCTG TTCAACGTC TGGTGGCTTC CATCGTTATC CTGCTGGCCA CGCCCATCGC 3540
 CGCCATCATC ATCCCGCCCC TTGGCCTCAT CTACTTCTTC GTCCAGAGGT TCTACGTGGC 3600
 25 TTCTTCCCG CAGCTGAAGC GCCTCGAGTC GGTCAAGCCG TCCCGGTCT ATTCCCATTT 3660
 CAACGAGACC TTGCTGGGGG TCAGCGTCAT TCGAGCCTTC GAGGAGCAGG AGCGCTTCAT 3720
 CCACCAGAGT GACCTGAAGG TGGACGAGAA CCAGAAGGCC TATTACCCCA GCATCGTGGC 3780
 AACACAGTGG CTCGCGGTGC GGTGGAGTG TGTGGGCAAC TGCATCGTTC TGTGCTGCTG 3840
 CCTGTTTGGC GTGATCTCCA GGCACAGCCT CAGTGTCTGC TTGGTGGGCC TCTCAGTGTG 3900
 30 TTACTCATTG CAGGTACCCA CGTACTGAA CTGGCTGGTT CGGATGTCAT CTGAATGGA 3960
 AACCAACATC GTGGCCGTGG AGAGGCTCAA GGAGTATTCA GAGACTGAGA AGGAGCGGCC 4020
 CTGGCAAAAT CAGGAGACAG CTCGCCCCAG CAGCTGGCCC CAGGTGGGCC GAGTGGAAAT 4080
 CCGGAACCTG TGCTTGGCTG ACCGAGAGGA CCTGGACTTC GTTCTCAGGC ACATCAATGT 4140
 CACGATCAAT GGGGAGAGAA AGGTCCGCAT CGTGGGGCGG ACGGGAGCTG GGAAGTCGTC 4200
 35 CCTGACCCTG GGCCTATTTC GGATCAACGA GTCTGCCGAA GGAGAGATCA TCATCGATGG 4260
 CATCAACATC GCCAAGATCG GCCTGCACGA CCTCCGCTTC AAGATCACCA TCATCCCCCA 4320
 GGACCCGTGT TTGTTTTCGG GTTCCCCTCC AATGAACCTG GACCCATTCA GCCAGTACTC 4380
 GGAAGAAGAA GTCTGGAGCT CCTTGGAGCT GGCCACCTG AAGGACTTCG TGTGAGCCCT 4440
 TCCTGACAAG CTAGACCATG AATGTGCAGA AGCGGGGAG AACCTCAGTG TCGGGCAGCG 4500
 40 CCAGCTTGTG TGCTTAGCCC GGGCCCTGCT GAGGAAGACG AAGATCCTTG TGTGAGTGA 4560
 GGGCAGGCA GCCGTGAGCC TGGAAACGGA CGACCTCATC CAGTCCACCA TCCGGACACA 4620
 GTTCGAGGAC TGCAACGTCC TCACCATCGC CCACCGGCTC AACACCATCA TGGACTACAC 4680
 AAGGGTGTAT GTCTTGACA AAGGAGAAAT CCAGGAGTAC GGCGCCCCAT CGGACCTCCT 4740
 GCAGCAGAGA GGTCTTTTCT ACAGCATGGC CAAAGACGCC GGCTTGGTGT GAGCCCCAGA 4800
 45 CCTGGCATAT CTGGTCAGAA CTCGAGGGCC TATATGCCAG CGCCAGGGA GGAGTCAGTA 4860
 CCCTTGGTAA ACCAAGCCTC CCACACTGAA ACCAAACAT AAAAAACAAA CCGAGACAAC 4920
 CAAACATAT TCAAAGCAGC AGCCACGCC ATCCGGTCCC CTGCTGGAA CTGGCTGTGA 4980
 AGACCCAGGA GAGACAGAGA TGCGAACCA C

50 SEQ ID NO:22 QBH2 Protein sequence:
 Protein Accession #: AAB46618

1 11 21 31 41 51
 55 MALRGFCSD GSDPLWDWNV TWNTSNPDET KCFQNTVLVW VPCFYLWACF PFYFLYLSRH 60
 DRGYIQMTPL NKTKTALGFL LWIVCWADLF YSFWERSRGI FLAPVFLVSP TLLGITTLA 120
 TFLIQLERRK GVQSSGIMLT FWLVALVICAL AILRSKIMTA LKEDAQVDLF RDITFYVYFS 180
 LLLIQLVLSC FSDRSPFSE TIHDNPCEPE SSASFSLRIT FWITGLIVR GYRQPLEGSD 240
 60 LWSLNKEDTS EQVVPVLVKN WKKECAKTRK QPVKVVSYSK DPAQPKESSK VDANEVEEAL 300
 IVKSPQKEWN PSFLKVLKYK FGPYFLMSFF FKAHDLMMF SGFQILKLLI KPVNDTKAPD 360
 WQGYFYTVLL FVTACLQTLV LHQYFHICFV SGMRIKTAVI GAVYRKALVI TNSARKSSTV 420
 GEIVNLSMVD AQRFMDLATY INMIWSAPLQ VILALYLLWL NLGPSVLAVG AVMVLMVFPVN 480
 65 AVMAKTKTY QVAHMKSKDN RIKLMNEILN GIKVLKLYAW ELAFKDKVLA IRQEBLKVLK 540
 KSAYLSAVGT FTWVCTPFLV ALCTFAVYVT IDENNILDAQ TAFVSLALFN ILRFPNLILP 600
 MVISSIVQAS VSLKRLRIFL SHEELEPDSI ERFPVKDGGG TNSITVRNAT FTWASDPPT 660
 70 LNGITFSIPE GALVAVVGQV GCGKSSLLSA LLAEMDKVEG HVAIKGSVAY VPQQAQIOND 720
 SLRENILFGC QLEFPYYSRV IQACALLPDL EILPSGDRTE IGEKGVNLSG GQKQRVSLAR 780
 AVYSNADIYL FDDPLSAVDA HVGKHIFENV IGPKGMLKNK TRILVTHSMS YLPQVDVIV 840
 MSGGKISEMG SYQELLARDG AFAEFLRTYA STEQEQAEE NGVTGVSFGP KEAKQMENG 900
 75 LVTDSAGKQL QRQLSSSSSY SGDISRHHS TAELOKAEAK KEETWKLMEA DKAQTGQVKL 960
 SVYVDYMKAI GLFISFLSIF LFMCNHVSAL ASNYWLSLWT DDPIVNGTQE HTKVRLSVYG 1020
 ALGISQGIIV FGYSMAVSIG GILASRCLHV DLLHSILRSP MSFFERTPSG NLVNRFSKEL 1080
 DTVDSMIPEV IKMFMSGLFN VIGACIVILL ATPIAAIIIP PLGLIYFFVQ RPYVASSRQL 1140
 KRLESVSRSP VYSHFNETLL GVSIVRAFEE QERFIHQSDL KVDENQKAY PSIVANRWLA 1200
 80 VRLECVGNCI VLFPAALFAVI SRHSLSAGLV GLSVSYSLQV TTYLNLWLRM SSEMETNIVA 1260
 VERLKEYSET EKEAPVQIE TAPPSSWPQV GRVEFRNYCL RYREDLDFVL RHINVTINGG 1320
 EKVGIVGRGT AGKSSLTLGL FRINESAEGE IIDGINIAK IGLHDLRFKI TIIPQDFVL 1380
 SGLSRMNLDP FSQYSDEEVW TSELAHLKD FVSALPDKLD HECAEGGENL SVGQRQLVCL 1440
 ARALLRKTKI LVLDEATAAV DLETTDLIQS TIRTQFEDCT VLTIAHRLNT INDYTRVIVL 1500
 DKGEIQEYGA PSDLLQQRGL FYSMAKDAGL V

SEQ ID NO:23 PAA2 DNA SEQUENCE

Nucleic Acid Accession #: NM_013309

Coding sequence: 1-1290 (underlined sequences correspond to start and stop codons)

5

	1	11	21	31	41	51	
10	ATGGCCGGCT	CTGGCGCGTG	GAAGCGCCTC	AAATCTATGC	TAAGGAAGGA	TGATGCGCCG	60
	CTGTTTAA	ATGACACCAG	CGCTTTGAC	TTCTCGGATG	AGGCGGGGGA	CGAGGGGCTT	120
	TCTCGGTACA	ACAAACTTCG	AGTTGTGGTG	GCCGATGACG	GTTCGGAAGC	CCCGGAAAGG	180
	CCTGTTAACG	GGGCGCACCC	GACCTCCAG	GCCGACGATG	ATTCCTTACT	GGACCAAGAC	240
	TTACCTTTGA	CCAACAGTCA	GCTGAGTTTG	AAGGTGGACT	CCTGTGACAA	CTGCAGCAAA	300
15	CAGAGAGAGA	TACTGAAGCA	GAGAAAGGTG	AAAGCCAGGT	TGACCATTCG	TGCCGTTCTG	360
	TACTTGCTTT	TCATGATTGG	AGAACTTGTA	GGTGGATACA	TTGCAAAATAG	CCTAGCAATC	420
	ATGACAGATG	CACTTCATAT	GTTAACTGAC	CTAAGCGCCA	TCATACTCAC	CCTGCTTGCT	480
	TTGTGGCTAT	CATCAAAATC	ACCAACCAAA	AGATTACCTT	TTGGATTTC	TCGCTTAGAG	540
	GTTTTGTGAG	CTATGATTAG	TGTGCTGTTG	GTGTATATAC	TTATGGGATT	CCTCTTATAT	600
20	GAAGCTGTGC	AAAGAACTAT	CCATATGAAC	TATGAAATAA	ATGGAGATAT	AATGCTCATC	660
	ACCGCAGCTG	TTGGAGTTGC	AGTTAATGTA	ATAATGGGGT	TTCTGTTGAA	CCAGTCTGGT	720
	CACCGTCACT	CCCATPCCCA	CTCCCTGCC	TCAAATPCCC	CTACCAGAGG	TTCTGGGTGT	780
	GAACGTAACC	ATGGGCAGGA	TAGCCTGGCA	GTGAGAGCTG	CATTTGTACA	TGCTTTGGGA	840
	GATTTGTGAC	AGAGTGTGG	TGTGCTAATA	GCTGCATACA	TCATACGATT	CAAGCCAGAA	900
25	TACAAGATTG	CTGATCCCAT	CTGTACATAC	GTATTTTCAT	TACTTGTGGC	TTTACAACA	960
	TTTCGAATCA	TATGGGATAC	AGTAGTTATA	ATACTAGAAG	GTGTGCCAAG	CCATTGGAAT	1020
	GTAGACTATA	TCAAAGAAGC	CTGTATGAAA	ATAGAAGATG	TATATTCAGT	CGAAGATTTA	1080
	AATATCTGGT	CTCTCACTTC	AGGAAAATCT	ACTGCCATAG	TTACATACA	GCTAATTCCT	1140
	GGAAGTTTAT	CTAAATGGGA	GGAAGTACAG	TCCAAAGCAA	ACCATTATAT	ATTGAACACA	1200
30	TTTGGCATGT	ATAGATGTAC	TATTCAGCTT	CAGAGTTACA	GGCAAGAAGT	GGACAGAACT	1260
	TGTGCAAAAT	GTCAGAGTTC	TAGTCCCTGA				

SEQ ID NO:24 PAA2 Protein sequence:

Protein Accession #: NP_037441

35

	1	11	21	31	41	51	
40	MAGSGAWKRL	KSMRLKDDAP	LFLNDTSAPD	FSDEAGDEGL	SRFNKLRRVVV	ADDGSEAPER	60
	PVNGAHPQLQ	ADDDSLLDQD	LPLTNSQLSL	KVDSCDNCSK	QREILKQRKV	KARLTIAAVL	120
	YLLFMIGELV	GGYIANSIAI	MTDALHMLTD	LSAIIITLLA	LWLSSKSPTK	RFTFGFHRLE	180
	VLSAMISVLL	VYILMGFLLY	EAVQRTIHMN	YEINGDIMLI	TAAVGVAVNV	IMGFLNQS	240
	HRHSHSHSLP	SNFPTRGSGC	ERNHGDQSLA	VRAAFVHALG	DLVQSVGLI	AAYIIRFKPE	300
45	YKIADPICTY	VFSLLVAFVT	FRIIWDTVVI	ILEGVPSHLN	VDYIKEALMK	IEDVYSVEDL	360
	NIWLSLTSKKS	TAIVHQLIIP	GSSSKWEEVQ	SKANHLLLNT	FGMYRCTIQL	QSYRQEVDR	420
	CANQSSSP						

SEQ ID NO:25 PAA3 DNA SEQUENCE

Nucleic Acid Accession #: AB037765

Coding sequence: 375-2798 (underlined sequences correspond to start and stop codons)

50

	1	11	21	31	41	51	
55	GCCGAGTCGG	TGGCGGCTGC	AGGCTGGGAG	GGAGAAGTGC	TACGCCTTTG	CAGGTTGGCG	60
	AAGTGGTPTC	AGGCTACCCG	GCTAGTCTGG	CACGGCCCCG	TCTTCTGCCT	CCTCCTCCGT	120
	CGCTGGCGGG	CGGGAACGTG	TGGCCGCGCG	GCCTCGGGAA	CGGCCAGGT	CCCCGCCCGC	180
60	AGGTCCCGGG	CAGATAACAT	AGATCATCAG	TAGAAAACCT	CTTGAAGTTG	TTCAAGAAAA	240
	ATTTGAAAGT	AGCAAAATAG	AAATAAAGA	ATTAACAGCA	GATACAGAGG	ACAGCATGGA	300
	AGTGTGTGCT	TAGGAAACAG	AACACAGCAG	TGAAAAACA	GACAAAATCC	GCTCAGATAC	360
	AACTGCAGCT	GATAATGTTT	TCCGGCTTCA	ATGTCCTTAG	AGTTGGGATC	TCTTTTGTCA	420
	TAATGTGCAT	TTTTTACATG	CCAACAGTAA	ACTCTTTACC	AGAACTGAGT	CCTCAGAAAT	480
65	ATTTTAGTAC	ATTGCAACCA	GGTCTTGAAG	AACTGAATGA	GGCTGTTAGA	CCTCTGCAGG	540
	ACTATGGAAT	TTCAAGTGGC	AAGGTTAATT	GTGTCAAAGA	AGAAATATCA	AGATACTGTG	600
	GAAAAGAAAA	GGATTTGATG	AAAGCATATT	TATTTCAAGG	CAACATATTG	CTCAGAGAAT	660
	TCCCTACTGA	CACCTGTGTT	GATGTGAATG	CCATTGTGCG	CCATGTTCTC	TTTGCTCTTC	720
	TTTTTAGTGA	AGTGAATAT	ATTACCAACC	TGGAAGACCT	TCAGAACATA	GAAATGCTC	780
70	TGAAAGGAAA	AGCAAAATAT	ATATTCTCAT	ATGTAAGAGC	CATTGGAATA	CCAGAGCACA	840
	GAGCAGTCAT	GGAAGCCGGT	TTTGTGTATG	GGACTACATA	CCAATTTGTC	TTAACCACAG	900
	AAATTTGCCCT	TTTGGAAAGT	ATTGGCTCTG	AGGATGTGGA	ATATGCACAT	CTCTACTTTT	960
	TTCAATTGTA	ACTAGTCTTG	GACTTGACCC	AGCAATGTAG	AAGAACACATA	ATGGAACAGC	1020
	CATTGACTAC	ACTGAACATT	CACCTGTTTA	TTAAGACAAT	GAAAGCACCT	CTGTTGACTG	1080
75	AAGTTGCTGA	AGATCCTCAA	CAAGTTTCAA	CTGTCCATCT	CCAAGTGGGC	TTACCACTGG	1140
	TTTTTATTTG	TAGCCAACAG	GCTACTTATG	AAGCTGATAG	AAGAACTGCA	GAATGGGTTG	1200
	CTTGGCGTCT	TCTGGGAAAA	GCAGGAGTTC	TACTCTTGT	AAGGGACTCT	TTGGAAGTGA	1260
	ACATTCTCTA	AGATGCTAAT	GTGGTCTTCA	AAAGAGCAGA	AGAGGGAGTT	CCAGTGGAAAT	1320
	TTTTTGGTATT	ACATGATGTT	GATTTAATAA	TATCTCATGT	GGAAAATAAT	ATGCACATTG	1380
80	AGGAAATACA	AGAAGATGAA	GACAATGACA	TGGAAGGTCC	AGATATAGAT	GTTCAGGATG	1440
	ATGAAGTGGC	AGAAACTGTT	TTCAAGAGATA	GGAAGAGAAA	ATTACCTTTG	GAACCTACAG	1500

	TGGAACCTAAC	AGAAGAAACA	TTTAATGCAA	CAGTGATGGC	TTCTGACAGC	ATAGTACTCT	1560
	TCTATGCTGG	TTGGCAAGCA	GTATCCATGG	CATTTTTCGA	ATCCTATATT	GATGTGGCAG	1620
	TTAAACTGAA	AGGCACATCT	ACTATGCTTC	TTACTAGAAT	AAACTGTGCA	GATTGGTCTG	1680
5	ATGTATGTAC	TAAGCAAAAT	GTTACTGAAAT	TTCCATATCAT	AAAGATGTAC	AAGAAAGGCG	1740
	AGAACCAGAT	ATCTTATGCT	GGAAATGTTAG	GAACCAAGA	TCTCCTAAAA	TTTATCCAGC	1800
	TTCAACAGGAT	TTCATATCCA	GTGAATATAA	CATCGATCCA	AGAAGCAGAA	GAATATTTAA	1860
	GTGGGGAATT	ATATAAAGAC	CTCATCTTGT	ATTCTAGTGT	GTCAGTATTG	GGACTATTTA	1920
	GTCCAACCAT	GAAAAACAGCA	AAAGAAGATT	TTAGTGAAAGC	AGGAAACTAC	CTAAAGGAT	1980
10	ATGTTATCAC	TGGAATTTAT	TCTGAAGAAG	ATGTTTGTCT	ACTGTCAACC	AAATATGCTG	2040
	CAAGTCTTCC	AGCCCTGCTG	CTTGCCAGAC	ACACAGAAGG	CAAAATAGAG	AGCATCCAC	2100
	TAGCTAGCAC	ACATGCACAA	GACATAGTTC	AAATAATAAC	AGATGCACATA	CTGGAAATGT	2160
	TTCCGGAAAT	CACGTGGGAA	AATCTTCCCA	GTATTTTCAG	ACTTCAGAAA	CCATTATTGA	2220
	TTTTGTTCAG	TGATGGCACT	GTAAATCCCT	AATATAAAAA	AGCAATATTG	ACACTGGTAA	2280
15	AGCAGAAATA	CTTGATTCCA	TTTACTCCAT	GCTGGTTAAA	TCTAAAGAAT	ACTCCAGTGG	2340
	GGAGAGGAAT	CTTGCGGGCA	TATTTTGATC	CTCTGCCCTC	CCTTCCCTCT	CTTGTTTTGG	2400
	TGAATCTGCA	TTTCAAGTGC	CAAGTATTTG	CATTTCCCTC	AGACCAGGCT	ATAATTGAAG	2460
	AAAACTTGT	ATTGTGGCTG	AAGAAATTAG	AAGCAGGACT	AGAAAATCAT	ATCACAATTT	2520
	TACCTGTCTA	AGAAATGAAA	CCTCCTCTTC	CAGCTTATGA	TTTTCTAAGT	ATGATAGATG	2580
20	CCGCAACATC	TCAACGTGGC	ACTAGGAAAG	TTCCCAAGTG	TATGAAAGAA	ACAGATGTGC	2640
	AGGAGAATGA	TAAGGAACAA	CATGAAGATA	AATCGGCAGT	CAGAAAAGAA	CCGATTGAAA	2700
	CTCTGAGAA	AAAGCATTTG	AATAGAAGTA	ATTGGTTTAA	AGAAGCAGAA	AAATCATTTA	2760
	GACGTGATAA	AGAGTTAGGA	TGCTCAAAAG	TGAACATAAT	TTATAGGGCT	GTGGTTTCCA	2820
	AAATTTTTTT	GGCATGATAG	ACTTAATTTA	TTTCTTTAAA	GAATAATATT	AAATCATTTT	2880
25	AAGTTTGCAG	ACTAGTGCCA	TCCAATAGAA	TTATAATATA	AGTCACATAT	TTTATTTAAA	2940
	ATTTTCTAGT	AACTACATTA	AACAAAGTAA	AAGTGAGCAG	GGCAAAATAA	TTTTGATATT	3000
	ACTTTTCCAC	CAGTAGTATA	CCCAAAATAG	CGAAATATAG	AAATTTATTA	TGAGATATTT	3060
	TACATCTTTT	TTTGPACCAA	GTCTTCTAAA	TGCAGTACAT	ATTTTATACT	TACTGCAATT	3120
	CTTACTTCCG	AGTAGCCATA	TTTCAAGTGT	TCATTGCCAC	ATGTGGCCTG	TGACTACTGT	3180
30	ATTGGACAGT	TCAGTACTAG	ACAAAACTA	GCATAATTAA	CTTAGTCTTA	GCCATGATTT	3240
	CTATTGTGGT	TAAAAATAAA	CTCTAATCAC	AGTTAACTCC	ACAGTGCAAT	CATGCAGCTG	3300
	ACAGTTATAT	TTGTTTATTT	GGAGTCATGA	TATTAAAAATC	AGCGTTTGTC	AACCTCAGGG	3360
	GATATTTAGC	AATTGTCCGG	AGACATTTT	GATGTCATGA	CTAGGGCAGT	TATTGACATT	3420
	TAGTGAGTAG	AGGCCATGAG	TCCTGTCTAA	TAACTTGCAAT	TGGACAGCGC	CCCACAACAA	3480
35	AGAAATATCC	TGCCCGAAAT	GGTAGTCTGT	CCAAGGCTGA	GTAACCTTGT	GTTAAAAGTA	3540
	ACCTGTGGCA	CTGTAGGTTT	CCAGAAATTC	CTGGTTCTGC	TCACGTATCA	TGTTTGAATA	3600
	AATTTTGGCT	ATTAAAGATA	TGTATTAGAT	GGTCTTATCC	TGATTATTAC	CTGGATACAA	3660
	CTTGATCTTT	TCTAATATTT	TCAGAAAAGT	ATGGGATAAC	CCTAGAAGAG	GACTCAGAA	3720
	GATATTTATA	TTTTAAGTGA	GTCTTAAATC	CTCCTCTTAT	TTCTACAAAT	TATATGGCTA	3780
40	AATTTTCAAT	TGAACAGGGA	TTTCAAGTGT	TGCCATCTCC	TCATTGGAAG	AGAGGCTCCC	3840
	TCATCTGAAG	CGTCTCTGAA	ATCTACCTTT	GCAAGCTTCA	GACAAATCAG	TTGATCTCCC	3900
	TGAGCCACAC	GGCCTCATTC	TGTGAGGGAG	GGAAAGATTA	GCCAAAGAGT	TAATTTTCAT	3960
	TCCAAATCAC	TTAGCTGTGA	GACTGATCTG	TTTGTAGCAG	TTGTTTGTCT	CATTTTGTCT	4020
	CTGTGCATTT	TTTGAGACAT	TTTGTAGGAA	TATTCTATTT	GGTGCTCTAC	TGTATTTTTC	4080
45	TTTTTAATAT	CTACTTGATA	TCTTGTCTTT	TAAATTTTCT	TCACATATGG	TTTGCCTGAT	4140
	ACAACGTGAT	TTTATAACTG	AAATTTAAGG	AATCTAACAG	CTAAACTTCA	GTAAGTGCAT	4200
	MTATTTCCCT	ATAACATAGA	CCCGTTGCTA	CTCTCAGCAC	CCTCTCCTCA	ATTTTCTTTC	4260
	CTGTAGCATG	TGATCCCTCA	TTAAACTCAT	TTTCAATTGC	TTTTATTFTCT	AAATATGGGA	4320
	CAATGAGAGT	GAACCTTAAA	TATAGGTTGT	AGTAATAAAA	CATCATTAGC	CTAATTATTA	4380
50	GAAAAATGCTA	ATTAAGTACC	AGCACATAGA	AACATGAAAT	TGCTTAGTCA	TTGTACCTTT	4440
	GTCAGCAATT	TTGACAGTCA	TTAATGTTTG	TCATAATTTT	AAATAAAGTG	TCTGGGTTTC	4500
	AGAAATACCT	CAAAAAAATA	AAAAAA				

SEQ ID NO:26 PAA3 Protein sequence:

Protein Accession #: BAA92582

55	1	11	21	31	41	51	
	MFSGFNVRV	GISFVIMCIF	YMPVTNSLPE	LSPQKYFSTL	QPGLEELNEA	VRPLQDYGIS	60
	VAKVNCVKEE	ISRYCGKEKD	LMKAYLFKGN	ILLREFPDT	LFDVNAIVAH	VLFALLFSEV	120
60	KYITNLEDLQ	NIENALKGKA	NIIFSIVRAI	GIPEHRAVME	AGFVYGTYYQ	FVLTTETALL	180
	ESIGSEDVEY	AHLYFFHCKL	VLDLTQQCRR	TLMEQPLTTL	NIHLFIKTMK	APLLTEVAED	240
	PQQVSTVHLQ	LGLPLVFIVS	QQATYEADRR	TAENVAWRLL	GKAGVLLLLR	DSLEVNIPQD	300
	ANVVFKRAEE	GVPVEFLVLH	DVDLIISHVE	NNMHIEEIQE	DEDNDMEGPD	IDVQDDEVAE	360
65	TVFRDRKRKL	PLELTVELTE	ETFNATVMAS	DSIVLFYAGW	QAVSMFLQS	YIDVAVKLKG	420
	TSTMLLTRIN	CADWSDVCYK	QNVTEFPPIK	MYKKGENPVS	YAGMLGPKDL	LKFIQLNRIS	480
	YFVNITSIQE	ABEYLSGELY	KDLILYSSVS	VLGLFSPTMK	TAKEDFSEAG	NYLKGIVITG	540
	TYSEEDVLLL	STKYAASLPA	LLLARHTEGK	IESIPLASTH	AQDIVQIITD	ALLEMPPEIT	600
	VENLPSYFRL	QKPLILFSD	GTVPNPQYKA	ILTLVKQKYL	DSFTPCWNLN	KNTFVGRGIL	660
70	RAYFDLPLPL	PLLVLVNLHS	GGQVFAFSPD	QAIIEENLVL	WLKKLEAGLE	NHITILPAQE	720
	WKPLPLPAYDF	LSMIDAATSQ	RGTRKVPKCM	KETDVQENDK	EQHEDKSAVR	KEPIETLRIR	780
	HWNRSNWFKE	AEKSFRRDKE	LGCSSKVN				

SEQ ID NO:27 PAA5 DNA SEQUENCE

Nucleic Acid Accession #: NM_012449

Coding sequence: 66-1085 (underlined sequences correspond to start and stop codons)

75	1	11	21	31	41	51	
	CCGAGACTCA	CGGTCAAGCT	AAGGCGAAGA	GTGGGTGGCT	GAAGCCATAC	TATTTTATAG	60
80	AATTAATGGA	AAGCAGAAAA	GACATCAACA	ACCAAGAAGA	ACTTTGGAAA	ATGAAGCCTA	120

5
10
15
20

```

GGAGAAATTT AGAAGAAGAC GATTATTTTGC ATAAGGACAC GGGAGAGACC AGCATGCTAA 180
AAAGACCTGT GCTTTTGCAT TTGCACCAAA CAGCCCATGC TGATGAATTT GACTGCCCTT 240
CAGAACCTCA GCACACACAG GAACCTCTTC CACAGTGGCA CTGCGCAATT AAAATAGCTG 300
CTATTATAGC ATCTCTGACT TTTCTTTTACA CTCTCTGAG GGAAGTAATT CACCCTTTAG 360
CAACTTCCCA TCAACAATAT TTTTATAAAA TTCCAATCCT GGTCAATCAAC AAAGTCTTGC 420
CAATGGTTTC CATCACTCTC TTGGCATTGG TTTACCTGCC AGGTGTGATA GCAGCAATTG 480
TCCAACCTCA TAATGGGAACC AAGTATAAGA AGTTTCCACA TTGGTTGGAT AAGTGGATGT 540
TAACAAGAAA GCAGTTTGGG CTCTCTCAGT TC'TTTTTTGC TGACTGTCAT GCAATTTATA 600
GTCTGTCTTA CCCAATGAGG CGATCCTTACA GATACAAGTT GCTAAACTGG GCATATCAAC 660
AGGTCCAACA AAATAAAGAA GATGCCCTGA TTGAGCATGA TGT'TTGGAGA ATGGAGATTT 720
ATGTGTCTCT GGAATTTGTG GGATTTGGCA TACTGGCTCT GTTGGCTGTG ACATCTATTG 780
CATCTGTGAG TGACTCTTTG ACATGGAGAG AATTTCAC TAATTCAGAGC AAGCTAGGAA 840
TTGTTCCTCT TCTACTGGGC ACAATACACG CATTGATTTT TGCCTGGAAT AAGTGGATAG 900
ATATAAAACA ATTTGTATGG TATACACCTC CAACTTTTAT GATAGCTGTT TTCCTTCCAA 960
TTGTGTCTCT GATATTTAAG AGCATACTAT TCCTGCCATG CTTGAGGAAG AAGATACTGA 1020
AGATTAGACA TGTTTGGGAA GACGTCACCA AAATTAACAA AACTGAGATA TGTTCCCACT 1080
TGTAAGAATA CTGTTTACAC ACATTTTGTG TCAATATTGA TATATTTTAT CACCAACATT 1140
TCAAGTTTGT ATTTGTTAAT AAAATGATTA TTCAAGGAAA AAAAAAAAAA AAAAA

```

SEQ ID NO:28 PAA5 Protein sequence

Protein Accession #: NP_036581

25
30

```

1 11 21 31 41 51
MESRKDITNQ EELWKMPPRR NLEEDDYLHK DTGETSMLKR FVLLHLHQTA HADEFDCPSE 60
LQHTQELFPQ WHLPKIAAI IASLTFLYTL LREVIHPLAT SHQQYFYKIP ILVINKVLPM 120
VSTLLALVY LPGVIAAIQV LHNGTKYKKF PHWLDKWLMT RKQFGLLSFF FAVLHAIYSL 180
SYPMRRSYRY KLLNWAYQQV QONKEDAWIE HDVWRMEIYV SLGIVGLLAIL ALLAVTSIPS 240
VSDSLTWREF HYIQSKLGIV SLLLGTHIAL IFANWKWIDI KQFVWYTPPT FMAIVFLPIV 300
VLIFKSILFL PCLRKKILKI RHGWEDVTKI NKTEICSQL

```

SEQ ID NO:29 PAA7 DNA SEQUENCE

Nucleic Acid Accession #: NM_030774

Coding sequence: 1-963 (underlined sequences correspond to start and stop codons)

35
40
45
50
55
60
65
70
75
80

```

1 11 21 31 41 51
ATGAGTTCCT GCAACTTCAC ACATGCCACC TTTGTGCTTA TTGGTATCCC AGGATTAGAG 60
AAAGCCCAT TCTGGGTTGG CTTCCCCCTC CTTTCCATGT ATGTAGTGGC AATGTTTGGA 120
AACTGCATCG TGTCTTTCAT CGTAAGGACG GAACGCAGCC TGCACGCTCC GATGTACCTC 180
TTTCTCTGCA TGTCTTGAGC CATTGACCTG GCCTTATCCA CATCCACCAT GCCTAAGATC 240
CTTGCCCTTT TCTGGTTTGA TTCCCGAGAG ATTAGCTTTG AGGCCCTGCT TACCCAGATG 300
TTCTTTATTC ATGCCCTCTC AGCCATTGAA TCCACCATCC TGCTGGCCAT GGCCTTTGAC 360
CGTTATGTGG CCATCTTGCCA CCACTGCGC CATGCTGCAG TGCTCAACAA TACAGTAACA 420
GCCAGATTG GCATCGTGGC TGTGGTCCGC GGATCCCTCT TTTTTCCTCC ACTGCCTCTG 480
CTGATCAAGC GGTGGGCTT CTGCCACTCC AATGTCTCTC CGCACTCCTA TTGTGTCCAC 540
CAGGATGTAA TGAAGTTGGC CTATGCAGAC ACTTTGCCCA ATGTGGTATA TGGTCTTACT 600
GCCATTCTGC TGTGTCATGG CGTGGACGTA ATGTTTCATC CTTTGTCCCTA TTTTCTGATA 660
ATACGAACGG TTTGCAACT GCCTTCCAAG TCAGAGCGGG CCAAGGCCCT TGGAACTCTG 720
GTGTACACA TTGGTGTGGT ACTCGCCTTC TATGTGCCAC TTATTGGCCT CTCAGTGGTA 780
CACCCTTTG GAAACAGCCT TCATCCCAT TGTGCTGTTG TCATGGGTGA CATCTACCTG 840
CTGCTGCCTC CTGTGCTCAT TCCCATCATC TATGGTGCCA AAACCAACAA GATCAGAAAC 900
CGGTGCTGG CTATGTTCAA GATCAGCTGT GACAAAGACT TGCAGGCTGT GGGAGGCAAG 960
TGACCCTTAA CACTACACT CTCTTATCT TATTGGCTT GATAACATA ATTATTTCTA 1020
ACACTAGCTT ATTTCCAGTT GCCCATAAGC ACATCAGTAC TTTTCTCTGG CTGGAATAGT 1080
AAACTAAAGT ATGGTACATC TACCTAAAGG ACTATTATGT GGAATAATAC ATACTAATGA 1140
AGTATTACAT GATTTAAGA CTACAATAA ACCAAACATG CTTATAACAT TAAGAAAAAC 1200
AATAAAGATA CATGATTGAA ACCAAGTTGA AAAATAGCAT ATGCCCTTGA GGAATGTGTC 1260
TCAAATTACT AATGATTATG TGTGTGCTCT ACTTCTCTCT TCTTTTCTCT TCTTTTCTCT 1320
TTTATTATGG TTAGCTGTCA CATAACAATT TTTT'TTTTTT TGAGATGGGG TCTCGCTCTG 1380
TCACCAGGCT GGAGTGCAGT GGCAGCATCT CGGCTCACTG CAACCTCCAC ATCCCATGTT 1440
GAAGTAATTC TTCTGCCTCA GCCTCCCGAG TAGCTGGGAC TAGAGGAACG TGCCACCATG 1500
ACTGGCTAAT TTCTGTATT TTTTAGTAGA GACAGAGTTT CACCATGTTG GCCAGGATGG 1560
TCTCGATCTC CTGACCTTGT GATCCACCCG CCTCAGCCTC CCAAAGTGTG GGGATTACAG 1620
GTGTGAACCA CTGTGCCCGG CCTGTGTACA ACTTTT'TAAA TAGGGAATAT GATAGCTTCG 1680
CATGGTGGTG TGCACCTATA GCCCCCCTG CCTGGAAGAG TGAGGTGGGA GAATCGCTTG 1740
AGTCCAGGAG TTTTGGGTTA CAGTGATCCA CGATCGTACC ACTACACTCC AGCCTGGGCA 1800
ACAGAGCAAG ACCTGTCTC AAAGCATAAA ATGGAATAAC ATATCAAATG AAACAGGGAA 1860
AATGAAGCTG ACAATTTATG GAAGCCAGGG CTGTCTACAG TCTCTACTGT TATTATGCAT 1920
TACCTGGGAA TTTATATAAG CCCTTAATAA TARTGCCAAT GAACATCTCA TGTGTGCTCA 1980
CAATGTTCTG GCACATATAT AAGTGCTTCA CAGGTTTATG GTGTTCTTCG TAAC'TTATG 2040
GAGTAGGTAC CATTTGTGTC TCTTTATAT AAGTGAGAGA AATGAAGTTT ATATTATCAA 2100
GGGACTTAAA GTACACGGC TTGTGGGCAC TGTGCCAAGA TTTAAATTA AATTGATGG 2160
TTGAATACAG TTACTTAATG ACCATGTTAT ATTGCTTCCT GTGTAACATC TGCCATTTAT 2220
TTCTCAGCTG TACTCAATCC TCTGTTTCT CTCTGTTACA CACTAACATC AATGGCTTTG 2280
TACTTGTGAT GAGAGATAAC CTGCCCCTAG TTGTGGGCAA CACATGCAGA ATAATCCTGT 2340
TTTACAGCTG CPTTTCGTGA TCTTATGTCT TGCTTTTTC CAGATTCAGG GAGAATGTTG 2400
TTGTCTATTT GTCTCTTACA TCCTCTGAT CATGCTTCA TTTT'TAATG TGCTCTGTAC 2460
CTGTCAAAAA TTTTGAATGT ACACCACATG CTATTGTCTG AACTTGAGTA TAAGATAAAA 2520
TAAATTTTAA TTTTAAATTT T

```

SEQ ID NO:30 PAA7 PROTEIN SEQUENCE

Protein Accession #: NP_110401

5
1 11 21 31 41 51
| | | | | |
MSSCNFTHAT FVLIGIPGLE KAHFWVGFP L SMYVVMFNG NCIVVFIVRT ERS LHAPMYL 60
FLCMLAAIDL ALSTSTMPKI LALFWFDSRE ISFEACLTQM FFIHALSAIE STILLAMAFD 120
10 RYVAICHPLR HA AVLNNVT AQIGIVAVVR GS LFFFLPL LIKRLAFCHS NVLSHSYCVH 180
QDVMKLAYAD TLPNVVYGLT AILLVMGVDV MFISLSYFLI IRTVLQLPSK SERAKAFGTC 240
VSHIGVVLAF YVPLIGLSVV HRFGNSLHPI VRVVMGDIYL LLPVVINPII YGAKTKQIRT 300
RVLAMFKISC DKDLQAVGGK

SEQ ID NO:31 PAV6 DNA SEQUENCE

Nucleic Acid Accession #: XM_050837

Coding sequence: 1-1020 (underlined sequences correspond to start and stop codons)

20 1 11 21 31 41 51
| | | | | |
ATGAACTGGG AGCTGCTGCT GTGGCTGCTG GTGCTGTGCG CGCTGCTCCT GCTCTTGGTG 60
CAGCTGCTGC GCTTCCTGAG GGCTGACGGC GACCTGACGC TACTATGGGC CGAGTGGCAG 120
GGACGACGCC CAGAATGGGA GCTGACTGAT ATGGTGGTGT GGGTGACTGG AGCCTCGAGT 180
GGAATTGGTG AGGAGCTGGC TTACCACTTG TCTAACTAG GAGTTTCTCT TGTGCTGTCA 240
GCCAGAAGAG TGCATGAGCT GGAAGGGTG AAAAGAAGAT GCCTAGAGAA TGGCAATTTA 300
25 AAGCAAGGAA ATATACTGT TTTGCCCTT GACCTGACCG AACTGGTTC CCATGAAGCG 360
GCTACCAAG CTGTTCTCCA GGAGTTTGGT AGAATCGACA TTCTGGTCAA CAATGGTGA 420
ATGTCCAGC GTTCTCTGTG CATGGATACC AGCTTGGATG TCTACAGAAA GCTAATAGAG 480
CTTAATACT TAGGACCGGT GTCCTTGACA AAATGTGTTT TGCCTCACAT GATCGAGAGG 540
AAGCAAGGAA AGATTGTTAC TGTGAATAGC ATCCTGGGTA TCATATCTGT ACCTCTTTCC 600
30 ATTGGTACT GTGCTAGCAA GCATGCTCTC CGGGGTTTTT TTAATGGCCT TCGAACAGAA 660
CTTGCCACAT ACCCAGTAT AATAGTTTCT AACATTGACC CAGGACCTGT GCAATCAAT 720
ATTGTGGAGA ATTCCGTAGC TGGAGAAGTC ACAAAGACTA TAGGCAATAA TGGAGACCAG 780
TCCCAACAAG TGACAACACG TCGTTGTGTG CGGCTGATGT TAATCAGCAT GGCCAATGAT 840
TTGAAAGAAG TTTGGATCTC AGACAACCT TTCTTGTAG TAACATATTT GTGGCAATAC 900
35 ATGCCAACCT GGGCTGTGTG GATAACCAAC AAGATGGGGA AGAAAAGGAT TGAGAACTTT 960
AAGAGTGTG TGGATGCAGA CTCTTCTTAT TTTAAATCT TTAAGACAAA ACATGACTGA

SEQ ID NO:32 PAV6 Protein sequence

Protein Accession #: XP_050837

40 1 11 21 31 41 51
| | | | | |
MNWELLWL L V LCALLLLV QLLRFLRADG DLTLWAEWQ GRRPEWELTD MVVWVTGASS 60
GIGELAYQL SKLVSVLVS ARRVHELERV KRRCLENGNL KEKDILVLPL DLTDTGSHEA 120
45 ATKAVLQEFGRIDILVNNG MSQRSLCMDT SLDVYRKLIE LNYLGTVSLT KCVLPHEMER 180
KQGLIVTVNS ILGLISVPLS IGVCAKHAL RGFENGLRTE LATVPGLIIV NICFPVQSN 240
IVENSLAGEV TKTTGNNGDQ SHKMTTSRCV RLMLISMND LKEVWISQEP FLVTVYLWQY 300
MPTWAWWIIN KMGKKRIENF KSGVDADSSY FKIFKTKHD

SEQ ID NO:33 PBA6 DNA SEQUENCE

Nucleic Acid Accession #: NM_006853

Coding sequence: 26-874 (underlined sequences correspond to start and stop codons)

55 1 11 21 31 41 51
| | | | | |
AGGAATCTGC GCTCGGGTTC CGCAGATGCA GAGGTTGAGG TGGCTGCGGG ACTGGAAGTC 60
ATCGGGCAGA GGTCTCACAG CAGCCAAGGA ACCTGGGGCC CGCTCCTCCC CCTCCAGGC 120
CATGAGGATT CTGCAGTTAA TCCTGCTTGC TCTGGCAACA GGGCTTGTAG GGGGAGAGAC 180
60 CAGGATCATC AAGGGGTTCG AGTGCAAGCC TCACTCCAG CCCTGGCAGG CAGCCCTGTT 240
CGAGAAGACG CGGCTACTCT GTGGGGCGAC GCTCATCGCC CCCAGATGGC TCCTGACAGC 300
AGCCCACTGC CTCAGCCCC GCTACATAGT TCACCTGGGG CAGCACAAACC TCCAGAAGGA 360
GGAGGGCTGT GAGCAGACCC GGACAGCCAC TGAGTCCTTC CCCACCCCG GCTTCAACAA 420
CAGCCTCCCC AACAAAGACC ACCGCAATGA CATCATGCTG GTGAAGATGG CATCGCCAGT 480
65 CTCCATCACC TGGGCTGTGC GACCCCTCAC CCTCTCTCA CGCTGTGTCA CTGCTGGCAC 540
CAGCTGCCTC ATTTCCGGCT GGGGCAGCAC GTCCAGCCCC CAGTTACGCC TGCCTCACAC 600
CTTGCGATGC GCCAACATCA CCATCATTGA GCACCAGAAG TGTGAGAACG CCTACCCCGG 660
CAACATCACA GACACCATGG TGTGTGCCAG CGTGCAGGAA GGGGGCAAGG ACTCTGCCA 720
GGGTGACTCC GGGGGCCCTC TGGTCTGTAA CCAGTCTCTT CAAGGCATTA TCTCTGGGG 780
70 CCAGGATCCG TGTGCGATCA CCCGAAAGCC TGGTGTCTAC ACGAAAAGTCT GCAAAATATG 840
GGACTGGATC CAGGAGACGA TGAAGAACAA TTAGACTGGA CCCACCCACC ACAGCCCATC 900
ACCTTCCATT TCCACTTGGT GTTTGGTTCC TGTTCACCTT GTTAATAAGA AACCTTAAGC 960
CAAGACCCTC TACGAACATT CTTTGGGCCCT CTTGGACTAC AGGAGATGCT GTCACTTAAT 1020
AATCAACCTG GGGTTCGAAA TCAGTGAGAC CTGGATTCAA ATTCTGCCTT GAAATATGTT 1080
75 GACTCTGGGA ATGACAACAC CTGGTTTGT CTCTGTGTGA TCCCCAGCCC CAAAGACAGC 1140
TCCTGGCCAT ATATCAAGGT TTCAATAAAT ATTGCTATAA TGAGTG

SEQ ID NO:34 PBA6 PROTEIN SEQUENCE

Protein Accession #: NP_006844

1	11	21	31	41	51	
MRILQLILLA	LATGLVGET	RIIKGFCKP	HSQFWQAALF	EKTRLLCGAT	LIAPRWLLTA	60
AHCLKPRYIV	HLGQHNLOKE	EGCEQTRTAT	ESFPHPGFNN	SLPNKDHRND	IMLVKMASPV	120
SITWAVRPLT	LSSRCVDTAGT	SCLISGWGST	SSPQLRLPHT	LRCANITIE	HQKCNAYPG	180
NITDTMVCAS	VQEGGKDSQC	GDSGGPLVCN	QSLQGIISWG	QDFCAITRKP	GVYTKVCKYV	240
DWIQETMKNN						

SEQ ID NO:35 PBC1 DNA SEQUENCE

Nucleic Acid Accession #: NM_001775

Coding sequence: 70-972 (underlined sequences correspond to start and stop codons)

1	11	21	31	41	51	
CTAAAGCTCT	CTTGCTGCCT	AGCCTCCTGC	CGGCTCATC	TTGCCCCAGC	CAACCCCGCC	60
TGGAGCCCTA	TGGCCAACCTG	CGAGTTCCAGC	CCGGTGTCCG	GGGACAAACC	CTGCTGCCGG	120
CTCTCTAGGA	GAGCCCAACT	CTGTCTTGCG	GTCAATATCC	TGGTCTCGAT	CCTCGTCGTG	180
GTGCTCGCGG	TGGTCGTCCC	GAGGTGGCGC	CAGACGTGGA	CGGGTCCGGG	CACCACCAAG	240
CGCTTTCCCG	AGACCGTCTC	GGCGCGATGC	GTCAAGTACA	CTGAAATTC	TCCTGAGATG	300
AGACATGTAG	ACTGCCAAGG	TGTATGGGAT	GCTTTCAAGG	GTGCATTTAT	TTCAAAACAT	360
CCTTGCAACA	TTACTGAAGA	AGACTATCAG	CCACTAATGA	AGTTGGGAAC	TCAGACCGTA	420
CCTTGCAACA	AGATTCTTCT	TTGGAGCAGA	ATAAAAGATC	TGGCCCATCA	GTTCACACAG	480
GTCCAGCGGG	ACATGTTTAC	CCTGGAGGAC	ACGCTGCTAG	GCTACCTTGC	TGATGACCTC	540
ACATGGTGTG	GTGAATTCAA	CACITCCAAA	ATAAATATC	AATCTTGCCT	AGACTGGAGA	600
AAGGACTGCA	GCAACAACCC	TGTTTCAGTA	TTCTGGAAAA	CGGTTTCCCG	CAGGTTTGCA	660
GAGGCTGCCT	GTGATGTGGT	CCATGTGATG	CTCAATGGAT	CCCGCAGTAA	AATCTTTGAC	720
AAAACAGCAG	CTTTTGGGAG	TGTGGAAGTC	CATAATTTCG	AACCAGAGAA	GGTTCAGACA	780
CTAGAGGCTT	GCTGTGATACA	TGGTGGAAAG	GAAGATTCCA	GAGACTTATG	CCAGGATCCC	840
ACCATAAAG	AGCTGGAATC	GATTATAAGC	AAAAGGAATA	TTCAATTTTC	CTGCAAGAAT	900
ATCTACAGAC	CTGACAAAGT	TCTTCAGTGT	GTGAAAAATC	CTGAGGATTC	ATCTTGCACA	960
TCTGAGATCT	GAGCCAGTGC	CTGTGGTTGT	TTTAGCTCCT	TGACTCCTTG	TGGTTTATGT	1020
CATCATACAT	GACTCAGCAT	ACCTGCTGGT	GCAGAGCTGA	AGATTTTGGA	GGGTCTCTCA	1080
CAATAAGGTC	AATGCCAGAT	ACGGAAGCCT	TTTTCCTCAA	AGTCTTAAAA	TAACCTTATAT	1140
CATCAGCATA	CCTTTATGTG	GATCTATCAA	TAGTCAAGAA	AAATTATTGT	ATAAGATTAG	1200
AATGAAAATT	GTATGTTAAG	TTACTTCCTT	TAG			

SEQ ID NO:36 PBC1 Protein sequence

Protein Accession #: NP_001766

1	11	21	31	41	51	
MANCEFSFVS	GDKPCCRLSR	RAQLCLGVSI	LVLILVVVLA	VVVPWRQWTW	SGPGTTKRFP	60
ETVLARCVKY	TEHPHMRHR	DCQSVWDAFK	GAFISKHPNC	ITEEDYQPLM	KLGTQTVPCN	120
KILLWSRIKD	LAHQFTQVQR	DMFTLEDTL	GYLADDLTWC	GEFNTSKINY	QSCPDRWKDC	180
SNNPVSFVFK	TVSRRAFEAA	CDVVHVMLNG	SRSKIFDKNS	TFGSVEVHNL	QPEKVQTLA	240
WVIHGGREDS	RDLCQDPTIK	ELESII SKRN	IQFSCKNIYR	PKFLLQCVKN	PEDSSCTSEI	

SEQ ID NO:37 PBH1 DNA SEQUENCE

Nucleic Acid Accession #: XM_017718

Coding sequence: 1-3315 (underlined sequences correspond to start and stop codons)

1	11	21	31	41	51	
ATGTCTTTTC	GGGCAGCCAG	GCTCAGCATG	AGGAACAGAA	GGAATGACAC	TCTGGACAGC	60
ACCCGGACCC	TGTACTCCAG	CGCGTCTCGG	AGCACAGACT	TGTCTTACAG	TGAAAGCGAC	120
TTGGTGAATT	TTATTCAAGC	AAATTTTAAG	AAACGAGAAT	GTGTCTTCTT	TACCAAAGAT	180
TCCAAGGCCA	CGGAGAATGT	GTGCAAGTGT	GGCTATGCCC	AGAGCCAGCA	CATGGAAGGC	240
ACCCAGATCA	ACCAAAATGA	GAAATGGAAC	TACAAGAAAC	ACACCAAGGA	ATTTCTTACC	300
GACGCCCTTG	GGGATATTC	GTTTGAGACA	CTGGGGAAGA	AAGGGAAGTA	TATACGTCTG	360
TCTTGCGACA	CGGACGCGGA	AATCCTTTAC	GAGCTGCTGA	CCGAGCACTG	GCACCTGAAA	420
ACACCCAAAC	TGGTCAATTC	TGTGACCGGG	GGCGCCAAGA	ACTTCGCCCT	GAAGCCGCGC	480
ATGCGCAAGA	TCCTCAGCCG	GCTCATCTAC	ATCGCGCAGT	CCAAAGGTGC	TTGGATTCTC	540
ACGGGAGGCA	CCCATTATGG	CCTGATGAAG	TACATCGGGG	AGGTGGTGAG	AGATAACACC	600
ATCAGCAGGA	GTTTCAGAGG	GAATATTGTG	GCCATTGGCA	TAGCAGCTTG	GGGCAATGGT	660
TCCAACCGGG	ACACCCTCAT	CAGGAATTGC	GATGCTGAGG	GCTATTTTTC	AGCCCAAGTAC	720
CTTATGGATG	ACTTCACAAG	AGATCCACTG	TATATCTCTG	ACAACAACCA	CACACATTTC	780
TCTGCTCGTG	ACAATGGCTG	TCATGGACAT	CCCACGTGCG	AAGCAAGACT	CCGGAATCAG	840
CTAGAGAAGT	ATATCTCTGA	GCGCACTATT	CAAGATTCCA	ACTATGGTGG	CAAGATCCCC	900
ATTGTGTGTT	TTGCCCAAGG	AGGTGGAAAA	GAGACTTTGA	AAGCCATCAA	TACCTCCATC	960
AAAAATAAAA	TTCCCTTGTG	GGTGGTGGA	GGCTCGGGCC	AGATCGCTGA	TGTGATCGCT	1020
AGCCTGGTGG	AGGTGGAGGA	TGCCCTGACA	TCTTCTGCGG	TCAAGGAGAA	GCTGGTGCGC	1080
TTTTTACCCC	GCACGGTGTG	CCGGCTGCCT	GAGGAGGAGA	CTGAGAGTTG	GATCAAAATG	1140
CTCAAGAGAA	TTCTCGAATG	TTCTCACCTA	TTAACAGTTA	TTAAATGGAA	AGAAGCTGGG	1200
GATGAAATTG	TGAGCAATGC	CATCTCCTAC	GCTCTATACA	AAGCCTTCAG	CACCACTGAG	1260
CAAGACAAGG	ATAACTGGAA	TGGGCAGCTG	AAGCTTCTGC	TGGAGTGGAA	CCAGCTGGAC	1320
TTAGCCAATG	ATGAGATTTT	CACCAATGAC	CGCCGATGGG	AGTCTGCTGA	CCTTCAAGAA	1380
GTCATGTTTA	CGGCTCTTCAT	AAAGGACAGA	CCCAAGTTTG	TCCGCTCTTT	TCTGGAGAAT	1440
GGCTTGAACC	TACGGAAGTT	TCTCACCCAT	GATGTCTCTA	CTGAACCTCT	CTCCAACCC	1500
TTCAGCAGCG	TTGTGTACCG	GAATCTGCAG	ATCGCCAAGA	ATTCTATATA	TGATGCCCTC	1560

5
 10
 15
 20
 25
 30

```

CTCACGTTTG TCTGGAAGCT GGTTCGGAAC TTCCGAAGAG GCTTCCGGAA GGAAGACAGA 1620
AATGGCCGGG ACGAGATGGA CATAGAAGCT CACGACGTGT CTCTATTAC TCGGCACCCC 1680
CTGCAAGCTC TCTTCATCTG GGCCATTCTT CAGAATAAGA AGGAACCTCTC CAAAGTCATT 1740
TGGGAGCAGA CCAAGGGGCTG CACTCTGGCA GCCCTGGGAG CCAGCAAGCT TCTGAAGACT 1800
CTGGCCAAAG TGAAGAACGA CATCAATGCT GCTGGGGAGT CCGAGGAGCT GGCATAATGAG 1860
TACGAGACCC GGGCTGTTGA GCTGTTCACT GAGTGTACA GCAGCGATGA AGACTTGGCA 1920
GAACAGCTGC TGGTCTATTC CTGTGAAGCT TGGGGTGGAA GCAACTGTCT GGAGCTGGCG 1980
GTGGAGGCCA CAGACCAGCA TTTCATCGCC CAGCCTGGGG TCCAGAATTT TCTTCTAAG 2040
CAATGGTATG GAGAGATTTC CCGAGACACC AAGAACTGGA AGATTATCCT GTGTCTGTTT 2100
ATTATACCTT TGGTGGGCTG TGGCTTTGTA TCATTTAGGA AGAAACCTGT CGACAAGCAC 2160
AAGAAGCTGC TTTGGTACTA TGTGGCGTTC TTCACCTCCC CCTTCGTGGT CTTCTCCTGG 2220
AATGTGGTCT TCTACATCGC CTTCCTCCTG CTGTTTGCTT ACGTGTCTGT CATGGATTTC 2280
CATTCGGTGC CACACCCCCC CGAGCTGGTC CTGTACTCGC TGGTCTTTGT CCTCTCTGT 2340
GATGAAGTGA GACAGTGGTA CGTAAATGGG GTGAATTATT TTAAGTACCT GTGGAATGTG 2400
ATGGACACGC TGGGGCTTTT TTAATTCATA GCAGGAATTG TATTTCGGCT CCACTCTTCT 2460
AATAAAGCTC CTTTGTATTC TGGACGAGTC ATTTTCTGTC TGGACTACAT TATTTTCACT 2520
CTAAGATTGA TCCACATTTT TACTGTAAGC AGAAACTTAG GACCCAAGAT TATAATGCTG 2580
CAGAGGATGC TGATCGATGT GTTCTTCTTC CTGTTCCTCT TTGCGGTGTG GATGGTGGCC 2640
TTTGGCGTGG CCAGGCAAGG GATCCTTAGG CAGAATGAGC AGCGCTGGAG GTGGATATTC 2700
CGTTCGGTCA TCTACGAGCC CTACCTGGCC ATGTTTCGGC AGGTGCCAGC TGACGTGGAT 2760
GGTACCAGCT ATGACTTTGC CCACTGCACC TTCACTGGGA ATGAGTCCAA GCCACTGTGT 2820
GTGGAGCTGG ATGAGCACAA CCTGCCCGG TCCCGCGAGT GGATCACCAT CCCCTGGTG 2880
TGCATCTACA TGTATCCAC CAACATCCTG CTGTGCAACC TGCTGGTCCG CATGTTTGGC 2940
TACACGGTGG GCACCGTCCA GGAGAACAAT GACCAGGTCT GGAAGTTCCA GAGGTACTTC 3000
CTGGTGCAGG AGTACTGCAG CCGCTCAAT ATCCCTTCC CCTTCATCGT CTTCTGCTAC 3060
TCTTACATGG TGGTGAAGAA GTGCTTCAAG TGTGTGTCGA AGGAGAAAAA CATGGAGTCT 3120
TCTGTCTGCT GTTCTCAAAA TGAAGACAAT GAGACTCTGG CATGGGAGGG TGTCATGAAG 3180
GAAAACTACC TTGTCAAGAT CAACACAAA GCCAACGACA CCTCAGAGGA AATGAGGCAT 3240
CGATTTAGAC AACTGGATAC AAAGCTTAAT GATCTCAAGG GTCTTCTGAA AGAGATTGCT 3300
AATAAATCA AATGA
  
```

SEQ ID NO:38 PBH1 Protein sequence

Protein Accession #: XP_017718

35
 40
 45
 50
 55

```

1 11 21 31 41 51
| | | | |
MSFRAARLSM RNRNDTLDLST TRLYSASR STDLSYSED LVNFIQANFK KRECVFTKD 60
SKATENVCKC GYAQSQHMEG TQINQSEKWN YKKHTKEFPT DAFGDIQFET LGKKGYIRL 120
SCDTPDAELLY ELLTQHWHLK TPNLVISVTG GAKNFALKPR MRKIFSRLIY IAQSKGAWIL 180
TGGTHYGLMK YIGEVVDNT ISRSSEENIV AIGIAAWGMV SNRDTLIRNC DAEGYFLAQY 240
LMDDFTRDPL YILDNNHTHL LLDVNGCHGH PTVEAKLRNQ LEKYISERTI QDSNYGGKIP 300
IVCAQGGGK ETLKAINTSI KNKIPCVVVE GSGQIADVIA SLVEVEDALT SSAVKEKLV 360
FLPRTVSRLP EEETESWIKW LKEILECSHL LTVIKMEEAG DEIVSNALISY ALYKAFSTSE 420
QDKDNWNGQL KLLLEWQOLD LANDEIFTND RRWESADLQE VMFTALIKDR PKFVRLFLEN 480
GLNLRKFLTH DVLTELSFNH FSTLVYRNLO IAKNSYNDAL LTFVWKLVAN FRRGFRKEDR 540
NGRDEMDEL HDVSPITRHP LQALFIWAIL QNKKELSKVI WEQTRGCTLA ALGASKLLKT 600
LAKVKNDINA AGESEELANE YETRAVELFT ECVSSDEDLA EQLLVYSCEA WGSNCLELA 660
VEATDQHFIA QPGVONFLSK QWYGEISRDT KNWKIILCLF IIPLVGCGFV SFRKKPVDKH 720
KKLLWYVAF FTSFVVFVSW NVVFYIAFLL LFAYVLLMDF HSPVHPPELV LYSLVFVLC 780
DEVQRWYVNG VNYFTDLWNV MDTLGLFYFI AGIVFRLHSS NKSSLYSGRV IFCLDYIIFT 840
LRLIHIFTVS RNLGPKIIML QRMLIDVFFF LFLFAVWMA FGVARQILR QNEQRWRWIF 900
RSVYIEPYLA MFGQVPSVDV GTTYDFAHCT FTGNESKPLC VELDEHNLP FPEWITIPLV 960
CIYMLSTNLI LVNLLVAMFG YTVGTVQENN DQVWKQRYF LVQEYCSRLN IPFFFIYFAY 1020
FYMVVKCKFK CCCKEKNMES SVCCFKNEDN ETLAWEGVMK ENYLVKINTK ANDTSEEMRH 1080
RFRQLDTKLN DLKGLLKEIA NKIK
  
```

SEQ ID NO:39 PBH3 DNA SEQUENCE

Nucleic Acid Accession #: XM_011804

Coding sequence: 1-558 (underlined sequences correspond to start and stop codons)

60
 65
 70

```

1 11 21 31 41 51 -
| | | | |
ATGCTCTGCC TGTCTCTGTT CCACCTGCTA GAATTCTGTT TACTACTGAA CCAATTTTCC 60
AGAGCAGTCG CGGCCAAATG GAAGGACGAT GTTATTAAAT TATGCGGCGC CGAATTAGTT 120
CGCGCGCAGA TTGCCATTG CGGCATGAGC ACCTGGAGCA AAAGGTCTCT GAGCCAGGAA 180
GATGCTCCTC AGACACCTAG ACCAGTGGCA GAAATGTGAC CATCCTTCAT CAACAAAGAT 240
ACAGAACTA TAATTATCAT GTTGAATTC ATTGCTAATT TGCCACCGGA GCTGAAGGCA 300
GCCCTATCTG AGAGGCAACC ATCATTACCA GAGCTACAGC AGTATGTACC TGCAATAAAG 360
GATTCCAAT TTAGCTTTGA AGAATTAAAG AAACCTATTC GCAATAGGCA AAGTGAAGCC 420
GCAGACAGCA ATCCTTCAGA ATTAATAATC TTAGGCTTGG ATACTCATTC TCAAAAAAAG 480
AGACGACCC TACGTGCACT GTTTGAGAAA TGTTCCTTAA TTGGTTGTAC CAAAAGGCTC 540
CTTGTCTAAT ATTCCTGA
  
```

SEQ ID NO:40 PBH3 PROTEIN SEQUENCE

Protein Accession #: NP_008842

75

```

1 11 21 31 41 51
| | | | |
MPRLFLFHLH EFCLLLNQFS RAVAAKWKDD VIKLCGRELV RAQIAICGMS TWSKRSLSQE 60
  
```


DAPQTPRPVA EIVPSFINKD TETIIIMLEF IANLPPELKA ALSERQPSLP ELQQYVPALK 120
 DSNLSFEEFK KLIRNRQSEA ADSNPSELKY LGLDTHSQKK RRPYVALFEK CCLIGCTKRS 180
 LAKYC

5

SEQ ID NO:41 PBH5 DNA SEQUENCE

Nucleic Acid Accession #: NM_005845

Coding sequence: 1-3978 (underlined sequences correspond to start and stop codons)

10 1 11 21 31 41 51
 | | | | | |
 ATGCTGCCCG TGTACCAGGA GGTGAAGCCC AACCCGCTGC AGGACGCGAA CCTCTGCTCA 60
 CGCGTGTTC TCTGGTGGCT CAATCCCTTG TTAAAAATG GCCATAAACG GAGATTAGAG 120
 GAAGATGATA TGTATTACAGT GCTGCCAGAA GACCGCTCAC AGCACCTTGG AGAGGAGTTG 180
 CAAGGGTTCT GGGATAAAGA AGTTTAAAGA GCTGAGAATG ACGCACAGAA GCCTTCTTTA 240
 15 ACAAGAGCAA TCATAAAGT TTACTGGAAA TCTTATTTAG TTTTGGGAAT TTTTACGTTA 300
 ATTGAGGAAA GTGCCAAAGT AATCCAGCCC ATATTTTGGG GAAAAATTAT TAATTATTTT 360
 GAAAAATTATG ATCCCATGGA TTCTGTGGCT TTGAACACAG CGTACGCCCTA TGCCACGGTG 420
 CTGACTTTT GCACGCTCAT TTTGGGTATA CTGCATCACT TATATTTTTA TCACGTTTCAG 480
 TGTGCTGGGA TGAGGTACG AGTAGCCATG TGCCATATGA TTTATCGGAA GGCACCTCGT 540
 20 CTTAGTAACA TGGCCATGGG GAAGACAACC ACAGGCCAGA TAGTCAATCT GCTGTCCAAT 600
 GATGTGAACA AGTTTGATCA GGTGACAGTG TTCTTACACT TCCTGTGGGC AGGACCACATG 660
 CAGGCGATCG CAGTGACTCG CCTACTCTGG ATGGAGATAG GAATATCTGT CCTTGTCTGG 720
 ATGGCAGTTC TAATCATTTCT CCTGCCCTTG CAAAGCTGTT TTGGGAAGTT GTTCTCATCA 780
 CTGAGGAGTA AAAGTCAAC TTTCACGGAT GCCAGGATCA GGACCATGAA TGAAGTTATA 840
 25 ATGGTATAAA GATAATAAAA AATGTACGCC TGGGAAAAGT CATTTTCAAA TCTTATTACC 900
 AATTTGAGAA AGAAGGAGAT TTCCAAGATT CTGAGAAATT CCTGCCTCAG GGGGATGAAT 960
 TTGGCTTCGT TTTTCAGTGC AAGCAAAATC ATCGTGTTTG TGACCTTCAC CACCTACGTG 1020
 CTCTCGGCA GTGTGATCAC AGCCAGCCGC GTGTTTCGTGG CAGTGACGCT GTATGGGGCT 1080
 GTGCGGCTGA CGGTTACCTC CTCTTCCCTC TCAGCCATG AGAGGGTGTG AGAGGCAATC 1140
 30 GTACAGCATCC GAAGAATCCA GACCTTTTTC CTACTTGATG AGATATCACA GCGCAACCGT 1200
 CAGCTGCCGT CAGATGGTAA AAAGATGGTG CATGTGCAGG ATTTTACTGC TTTTGGGGAT 1260
 AAGGCATCAG AGACCCCAAC TCTACAAGGC CTTTCTCTTA CTGTCAGACC TGGCGAATTG 1320
 TTAGCTGTGG TCGGCCCTCAG AGCCAGCCGC AAGTCATCAC TGTAAAGTGC CGTGTCTGGG 1380
 GAATTGGCCC CAAGTACCGG GCTGGTCAGC GTGCATGGAA GAATTGCCCTA TGTGTCTCAG 1440
 35 CAGCCCTGGG TGTCTCTGGG AACTCTGAGG AGTAATATTT TATTTGGGAA GAAATACGAA 1500
 AAGGAACGAT ATGAAAAAGT CATAAAGGCT TGTGCTCTGA AAAAGGATTT ACAGCTGTGT 1560
 GAGGATGGTG ATCTGACTGT GATAGGAGAT CGGGGAACCA CGCTGAGTGG AGGGCAGAAA 1620
 GCACGGGTAA ACCTTGCAAG AGCAGTGTAT CAAGATGCTG ACATCTATCT CCTGGACGAT 1680
 40 CCTCTCAGTG CAGTAGATGC GGAAGTTAGC AGACACTTGT TCGAACTGTG TATTTGTCAA 1740
 ATTTTGCATG AGAAGATCAC AATTTTAGTG ACTCATCAGT TGCAGTACCT CAAAGCTGCA 1800
 AGTCAGATTC TGATATTGAA AGATGGTAAA ATGGTGACGA AGGGGACTTA CACTGAGTTC 1860
 CTAAATCTCG GTATAGATTG TGGCTCCCTT TTAAGAAGAG ATAATGAGGA AATGGAACAA 1920
 CCTCCAGTTC CAGGAATCTC CACACTAAGG AATCGTACCT TCTCAGAGTC TCTGGTTTGG 1980
 45 TCTCAACAAT CTCTAGACC CTCTTGAAA GATGGTGCTC TGGAGAGCCA AGATACAGAG 2040
 AATGTCCAGG TTACACTATC AGAGGAGAAC CGTCTGTAAG GAAAAGTTGG TTTTCAGGCC 2100
 TATAAGAATT ACTTTCAGAG TGGTGCTCAC TGGATTGTCT TCATTTTCTT TATTTCTCTA 2160
 AACACTGCAG CTCAGGTGTC CTATGTGCTT CAAGATTGGT GGCTTTCATA CTGGGCAAAC 2220
 AAACAAAGTA TGCTAAATGT CACTGTAAAT GGAGGAGGAA ATGTAACCGA GAAGCTAGAT 2280
 50 CTTAACTGGT TTATTCAGGT TTAAGTGTAG CTACCGTTCT TTTTGGCATA 2340
 GCAAGATCTC TATTTGTTAT CTACGTCCTT GTTAACTCTT CACAACTTT GCACAACAAA 2400
 ATGTTTGAGT CAATTTCTGAA AGCTCCGGTA TTATTTCTTG ATAGAAATCC AATAGGAAGA 2460
 ATTTTAAATC GTTTCPCAA AGACATTGGA CACTTGGATG ATTTGCTGCC GCTGACGTTT 2520
 TTAGATTTC TCCAGACATT GCTACAAGTG GTTGGTGTGG TCTCTGTGGC TGTGGCCGTG 2580
 55 ATTCCTTGGA TCGCAATACC CTTGGTTCCC CTTGGAATCA TTTTCATTTT TCTTCGGCGA 2640
 TATTTTGTGG AAACGTCAAG AGATGTGAAG CGCCTGGAAT CTACAACCTG GAGTCCAGTG 2700
 TTTTCCCACT GTTCATCTTC TCTCCAGGGG CTCTGGACCA TCCGGGCATA CAAAGCAGAA 2760
 GAGAGGTGTC AGGAACCTGT TGATGCACAC CAGGATTTAC ATTCAGAGGC TTGGTTCTTG 2820
 TTTTGTACAA CGTCCCGCTG GTTCGCGCTC CGTCTGGATG CCATCTGTGC CATGTTTGTG 2880
 60 ATCATGCTTG CCTTTGGGTC CCGTATCTG GCAAAAACTC TGGATGCCGG GCAGGTTGGT 2940
 TTGGCACTGT CCTATGCCCT CACGCTCATG GGGATGTTTC AGTGGTGTGT TCGACAAAGT 3000
 GCTGAAGTTG AGAATATGAT GATCTCAGTA GAAAGGGTCA TTGAATACAC AGACCTTGAA 3060
 AAAGAAGCAC CTTGGGAATA TCAGAAACGC CCACCACCAG CCTGGCCCCA TGAAGGAGTG 3120
 ATAATCTTTG ACAATGTGAA CTTCATGTAC AGTCCAGGTG GGCCTCTGGT ACTGAAGCAT 3180
 65 CTGACAGCAC TCATTAATAT ACAAGAAAAG GTTGGCATTG TGGGAAGAAC CGGAGCTGGA 3240
 AAAAGTTCCC TCATCTCAGC CCTTTTGTAGA TTGTCAGAAC CCGAAGGTAA AATTTGGATT 3300
 GATAAGATCT TGACAACCTG AATTTGGACT CACGATTTAA GGAAGAAAAT GTCAATCATA 3360
 CCTCAGGAAC CTGTTTGTGT CACTGGAACA ATGAGGAAAA ACCTGGATCC CTTTAATGAG 3420
 CACACGGATG AGGAACCTGT GAATGCCCTA CAAGAGGTAC AACTTAAAGA AACCATTGAA 3480
 70 GATCTTCTGT GTAAATGTTA TACTGAATTA GCAGAATCAG GATCCAATTT TAGTGTGGA 3540
 CAAAGACAAC TGGTGTGCCCT TGCCAGGGCA ATTTCTCAGGA AAAATCAGAT ATTGATTATT 3600
 GATGAAGCGA CGGCAATATG GGATCCAAGA ACTGATGAGT TAATACAAAA AAAATCCGG 3660
 GAGAAATTTG CCCACTGCAC CTGTCTAACC ATTGACACCA GATTGAACAC CATTATTGAC 3720
 AGCGACAAGA TAATGGTTTT AGATTCAAGG AGACTGAAAG AATATGATGA GCCGTATGTT 3780
 75 TTGCTGCAAA ATAAAGAGAG CCTATTTTAC AAGATGGTGC AACAACTGGG CAAGGCAGAA 3840
 CGCGCTGCCC TCACTGAAAC AGCAAAACAG GTATACCTCA AAAGAAAATA TCCACATATT 3900
 GGTCACACTG ACCACATGGT TACAAACACT TCCAATGGAC AGCCCTCGAC CTTAACTATT 3960
 TTCGAGACAG CACTGTGA

SEQ ID NO:42 PBH5 PROTEIN SEQUENCE

Protein Accession #: NP_005836

5	1	11	21	31	41	51	
	MLPVYQEVKP	NPLQDANLCS	RVFFWNLNPL	FKIGHKRRLE	EDDMYSVLPE	DRSQHLGEEL	60
	QGFWDKEVLR	AENDAQKPSL	TRAIKCYWK	SYLVLGIFTL	IEESAKVIQP	IFLGKIINYF	120
	ENYDPMDSVA	LNTAYAYATV	LTFTCLILAI	LHHLVYFHVQ	CAGMRLRVAM	CHMIYRKALR	180
10	LSNMAMGKTT	TGQIVNLLSN	DVNKFDQVTV	FLHFLWAGPL	QATAVTALLW	MEIGISCLAG	240
	MAVLIILLPL	QSCFGKLFSS	LRSKTATFTD	ARIRTMNEVI	TGIRIIRKMYA	WEKSFSNLIT	300
	NLRKKEISKI	LRSSCLRGMN	LASFFSASKI	IVFVTFTTYV	LLGSVITASR	VFVAVTLYGA	360
	VRLTVTLFFP	SAIERVSEAI	VSIIRRIQTF	LLDEISQRNR	QLPSDGKKMV	HVQDFTAFWD	420
	KASEPTTLQG	LSFTVRPGEL	LAVVGPVGAG	KSSLLSAVLG	ELAPSHGLVS	VHGRIAYVSQ	480
15	QPWVFSGTLR	SNILFGKYE	KERYEKVIKA	CALKKDLQLL	EDGDLTVIGD	RGTTLSGGQK	540
	ARVNLARAVY	QDADIYLLDD	PLSAVDAEVS	RHLFELCICQ	ILHEKITILV	THQLQYLKAA	600
	SQILILKDGK	MVQKGTYTEF	LKSGIDFGSL	LKKNDEESEQ	PPVPGTPTLR	NRTFSESSVW	660
	SQSSSRPSLK	DGALESQDTE	NVPVTLSEEN	RSEKVGFGQA	YKNYFRAGAH	WIVFIFLILL	720
	NTAAQVAVVL	QDWLWSYWAN	KQSMNLNVTN	GCGNVTEKLD	LNWYLGIIYS	LTVATVLFGI	780
20	ARSLLVFVYL	VNSSQTLHNK	MFESILKAPV	LFFDRNPGR	ILNRFSDKIG	HLDDLPLTF	840
	LDFTQLTLQV	VGVVSVAVAV	IPWIAIPLVP	LGIIIFILRR	YFLETSRDVK	RLESTTRSPV	900
	FSHLSSSLQG	LWTRAYKAE	ERCQELFDAH	QDLHSEAWFL	FLTTSRWFAV	RLDAICAMFV	960
	IIVAFGSLIL	AKTLDAGQVG	LALSYALTLM	GMFQWCVRQS	AEVENMMISV	ERVIEYTDLE	1020
	KEAPWEYQKR	PPPAWPHFEG	IIFDNVNFMY	SPGGPLVLKH	LTALIKSQEK	VGIVGRTGAG	1080
25	KSSILISALFR	LSEPEGKIWI	DKLLTTEIGL	HDLRKKMSII	PQEPVLTGT	MRKNLDPFNE	1140
	HTDEELWNAL	QEQVQKLTIE	DLPGKMDTEL	AESGSNFSVG	QRQLVCLARA	ILRKNQILII	1200
	DEATANVDPR	TDQLIQKKIR	EKFAHCTVLT	IAHRLNTIID	SDKIMVLDSG	RLKEYDEPVV	1260
	LLQNKESLFY	KMVQQLGKAE	AAALTETAKQ	VYFKRNYPHI	GHTDHMTVNT	SNGQPSTLTI	1320
	FETAL						

SEQ ID NO:43 PBQ7 DNA SEQUENCE

Nucleic Acid Accession #: NM_021233

Coding sequence: 34-1119 (underlined sequences correspond to start and stop codons)

35	1	11	21	31	41	51	
	ATGGGGAAG	TGTCCTGCTG	TGGCATGAAA	TAAATGAAAC	AGAAAATGAT	GGCAAGACTG	60
	CTAAGAACAT	CCTTTGCTTT	GCTCTTCTTT	GGCCTCTTTG	GGGTGCTGGG	GGCAGCAACA	120
	ATTTTCATGCA	GAAATGAGA	AGGGAAAGCT	GTGGACTGGT	TTACTTTTTA	TAAGTTACCT	180
40	AAAAGACAAA	ACAAGGAAAG	TGGAGAGACT	GGGTTAGAGT	ACCTGTACCT	AGACTCTACA	240
	ACTAGAAGCT	GGAGGAAGAG	TGAGCAACTA	ATGAATGACA	CCAAGAGTGT	TTTGGGAAGG	300
	ACATTACAA	AGCTATATGA	AGCATATGCC	TCTAAGAGTA	ACAACACAGC	CTATCTAATA	360
	TACAAATGATG	GAGTCCCTAA	ACCTGTGAAT	TACAGTAGAA	AGTATGGACA	CACCAAGGTT	420
	TTACTGCTGT	GGAACAGAGT	TCAAGGGTTC	TGGCTGATTC	ATTCCATCCC	TCAGTTTCCT	480
45	CCAATTCGGG	AAGAAGGCTA	TGATTATCCA	CCCACAGGGA	GACGAAATGG	ACAAAGTGGC	540
	ATCTGCATAA	CTTTCAGATA	CAACCAAGTA	GAGGCAATAG	ATTCTCAGCT	CTTGGTCTGC	600
	AACCCCAACG	CTTATAGCTG	CTCCATCCCA	GCCACCTTTC	ACCAGGAGCT	CATTACATG	660
	CCCCAGCTGT	GCACAGGGC	CAGCTCATCA	GAGATTCTCTG	GCAGGCTCCT	CACCACACTT	720
	CAGTCGGCCC	AGGGACAAAA	ATTCTCTCCAT	TTTGCAAAAGT	CGGATTCCTT	TCTTGACGAC	780
50	ATCTTTGTCAG	CCTGGATGGC	TCAACGGCTG	AAGACACACT	TGTTAACAGA	AACCTGGCAG	840
	CGAAAAAGAC	AAGAGCTTCC	TTCAAACTGC	TCCCTTCTCT	ACCATGTCTA	CAATATAAAA	900
	GCAATTAAAT	TATCACGACA	CTCTTATTTT	AGTCTCTATC	AAGATCACGC	CAAGTGGTGT	960
	ATTTCCCAAA	AGGGCACCAA	AAATCGCTGG	ACATGTATTG	GAGACCTAAA	TCGGAGTCCA	1020
	CACCAAGCCT	TCAGAAGCTG	AGGATTGATT	TGTACCCAGA	ATTGGCAAT	TTACCAAGCA	1080
55	TTTCAAGGAT	TAGTATTATA	CTATGAAAGC	TGTAAGTAAA	CTTGGTGAAT	GGACACAGGT	

SEQ ID NO:44 PBQ7 Protein sequence

Protein Accession #: NP_067056

60	1	11	21	31	41	51	
	MMARLLRTSF	ALLFLGLFGV	LGAATISCRN	EKGAVDWFT	FYKLPKRQNK	ESGETGLEYL	60
	YLDSTTRSWR	KSEQLMNDTK	SVLGRITLQL	YEAYASKSNN	TAYLIYNDGV	PKFVNYSRKY	120
	GHTKGLLLWN	RVQGFWLIHS	IPQFPPIPEE	GVDYPTGRR	NGQSGICITF	KYNQVEAIDS	180
65	QLLVCPNPNV	SCSIPATFHQ	ELIHPQLCT	RASSSEIPGR	LLTTLQSAQG	QKFLHFAKSD	240
	SFLDDIFAAW	MAQLKTHLL	TETWQRKRQE	LPSNCSLPYH	VYNIKAIKLS	RHSYFSSYQD	300
	HAKWCISQKG	TENRWTCIGD	LNRSPHQAFR	SGGFICTQNW	QIYQAFQGLV	LYYESCK	

SEQ ID NO:45 PCQ8 DNA SEQUENCE

Nucleic Acid Accession #: XM_030453

Coding sequence: 89-1273 (underlined sequences correspond to start and stop codons)

75	1	11	21	31	41	51	
	CGGTGCCCTG	GGGTGGAATA	TCCCCTACGA	ATTTAACCAA	GCGGACTTTA	ATGCCACTGT	60
	GAGTTTCATC	CAAAACCACT	TGGATGACAT	GGATGTCAAA	AAGGGTGTCT	CCTGGACCAC	120
	CATCCGCTAC	ATGATAGGAT	AGATTCAATA	TGGAGGCAGA	GTCACGTGAC	ACTATGATAA	180
	GAGATTGTTG	AACACATTTG	CTAAGGTTTG	GTTCAGTGAA	AATATGTTTG	GACCAGATTT	240
	CAGTTTTCAC	CAAGGATACA	ATATTCCAAA	ATGCAGCACA	GTGGATAACT	ATCTTCAGTA	300
	TATCCAGAGT	TTGCCTGCCT	ATGACAGCCC	TGAGGTGTTT	GGGCTGCACC	CCAATGCTGA	360

CATCACCTAC CAGAGCAAGC TGGCCAAGGA CGTGCTGGAC ACCATCCTAG GCATCCAACC 420
 CAAGGACACC TCTGGTGGAG GGGATGAGAC CCGGGAGGCG GTGGTGGCCC GGCTGGCTGA 480
 TGATATGCTG GAGAAGCTGC CCCAGACTA TGTCCCTTTT GAAGTAAAG AGAGGCTGCA 540
 5 GAAGATGGGG CCATTCCAGC CTATGAACAT TTTCCTCAGG CAGGAAATAG ACAGAATGCA 600
 AAGGGTACTC AGCCTTGTCC GCAGCACCCCT CACTGAGCTG AAACCTTGCTA TTGATGGCAC 660
 CATCATCATG AGCGAAAATC TGCAAGATGC ATTGGATTGC ATGTTTGATG CTAGAATCCC 720
 TGCTTGGTGG AAAAAAGCTT CTGGGTTTTT TAGTACACTG GGTTTCTGGT TTACTGAACT 780
 TATAGAAAGA AACAGCCAGT TTACCTCGTG GGTTTTCAAT GGCCGACCTC ACTGCTTTTG 840
 10 GATGACGGGT TTTTAAACC CCCAGGGATT TTAACTGCA ATGCGACAGG AAATAACTCG 900
 GGCCAACAAA GGCTGGGCTC TGGACAATAT GGTGCTTTGC AATGAAGTCA CCAATGGAT 960
 GAAGGACGAC ATTTCTACCC CTCCCACAGA GGGTGTCTAT GTCTATGGCT TATATCTTGA 1020
 AGGTGCTGGC TGGGACAAGA GGAACATGAA ACTCATTTGAA TCAAAGCCAA AAGTGCTCTT 1080
 TGAGTTGATG CCTGTATATA GGATTTATGC AGAAAACAAT ACTTTACGAG ATCCTCGGTT 1140
 15 TTACTCCTGT CCCATCTATA AGAAGCCAGT TCGAAGCGAC TTGAACCTACA TTGCCGCTGT 1200
 GGATCTCAGG ACAGCCACAGA CCCCTGAACA CTGGGTGCTC CGTGGGGTTG CCCTTCTGTG 1260
 TGATGTCAAG TAACATGTGG GGAGTGTCCC CACCCAATGC TTTGGAAAAT GCAAGATCTA 1320
 AATTAATGTA ACCTTTATTT CTGTATGACT GCTGGACAGT GTATGTTAGG TCGCTTATGC 1380
 AATTAATGAG CTGCATAGGT TTCCCTCACT CCTTAATTGG ATGCTTATAT TTTACTTGT 1440
 20 TCATCATTAG TGACCAATGT CTGAGTTTGT TGAAAATGTT ATTTAGTGAT ATAAAAGTAA 1500
 ATTTACAGCA TCCTAATGAA GTGTGGCCCT CAAATCCACA GTAGTATATT TTCTTCTTAC 1560
 TTCGCTCCGA AGACTGACTG TGATTTATAA AGCAAATATA TTTGCATGTG GACAAAGATT 1620
 AGATGGCAAG ATAGAAAAT AAGAACAGAT GTGATAGCAA GAATTATAGT TGGCTTGAAA 1680
 AAATGTGATG ATCAGAGAAA AAAATAAAAA AAGGGTAGAA ATATTAGACG GTGCGTAGGG 1740
 25 ACTTCTATG GACTTTTATT AATTAGGAAA CATTATCAAA GGAACCTTTC ACGTATTTT 1800
 CTTTAAATTC TGGTTAGATG TTATTAATAA TTCTTCATCT AACCTACTGA CTAGAAAATA 1860
 TAGTCAGTAC TAAATAGAAA TTGTGGTTTA TAAACTTTTG GTTAGCTCTG GATCTGTATA 1920
 ACTGCATTTT TTTGGATAAA CAGTTTGTGG TAGGTGGATA CCGGGAGACA AGTGTGGGTC 1980
 CCTCTCACTG GGCTTCATTC TGTGGACCAG GATCATTATT TCATGCTCAT GATCATGAGA 2040
 30 GTTAGGACTG AGTGGCTCCT GTGACTCCCA CCATCTTAGA TGATACTGTT TTCTTGTGAG 2100
 TTCTTTCTTT TGGTGTGGAT TAGTATATCA GTTGATTTGT GTGAATTGTG GTGAAACAAT 2160
 CATTTCAATT TGAAAAGCAA GTAATGAAAA TGTACGATC ATAGGAATTA ATAAAATGTT 2220
 TTTACTAAAA AAAAAAAAAA AAA

SEQ ID NO:46 PCQ8 Protein sequence

Protein Accession #: BAB15543

1 11 21 31 41 51
 MDVKKGVSWT TTRYMIGEIQ YGGRVTDVDD KRLNLTFAKV WFSENMFGPD FSYQGYNIP 60
 40 KCSTVDNYLQ YIQSLPAYDS PEVFLGHPNA DITYQSKLAK DVLDTILGIQ PKDTSGGGDE 120
 TREAVVARLA DDMLEKLPPD YVFFEVKERL QKMGPFQPMN IFLRQEIDRM QRVLSLVRST 180
 LTELKLAIDG TIIMSENLOQ ALDCMFDAIRI PAWWKKASWV FSTLGFWFTE LIERNSSQFTS 240
 WVENGRHCF WMTGFFNPQG FLTAMRQEIT RANKGWALDN MVLNCNEVTKW MKDDISTPPT 300
 45 EGVVYVGLYL EGAGWDKRNK KLIESKPKVL FELMPVIRIY AENNTLRDPR FYSCTPIYKKP 360
 VRTDLNYIAA VDLRTAQIPE HWVLRGVALL CDVK

SEQ ID NO:47 PDG5 DNA SEQUENCE

Nucleic Acid Accession #: AB033036

Coding sequence: 68-3349 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 GGAGCAGCCT ACAACTTCCAC AACCAGAAAC CACTACCCCT CAGGGGTTCG TTTCAGATAA 60
 55 AGATGACATG GGAAGGAGAA ATGCTGGCAT AGATTTCGGA TCCAGAAAAG CATCAGCAGC 120
 ACAGCCCATTA CCTGAAACA TGGACAATTC CATGGTTAGT GATCCACAAC CATACCATGA 180
 AGATGCAGCT TCTGGAGCTG AGAAGACAGA AGCCAGAGCT TCTCTCTCAC TGATGGTGGG 240
 AAGCCTTTCT ACAACCCAAG AGGAGGCCAT TCTCTCAGTA GCAGCAGAGG CTCAGGTGTT 300
 TATGAATCCT TCTCATATCC AGTTAGAAGA TCAAGAAGCT TTCAGCTTTG ATTTACAAAA 360
 60 GGCCCAATCC AAAATGGAGT CAGCCAGGGA TGTTCAAAC ATCTGCAAG AAAAGCCTTC 420
 TGGAAATGTT CACCAGACCT TTACAGCAAG TGTTTGGGT ATGACAAGTA CTACAGCCAA 480
 AGGAGATGTT TATGCCAAGA CTCTGCCTCC CAGAAGCCTT TTTCACTCCT CAAGGAAGCC 540
 TGATGCTGAA GAAGTCTCCT CAGATTTCAGA GAATATTCC T GAGGAGGGGG ATGGTTCTGA 600
 AGAACTGGCT CATGGTCACT CTTCCTCAGT CTTGGGGAAG TTTGAAGATG AACAAGAAGT 660
 65 CTCTCAGAAA TCAAAAAGTT TTGTTGAGGA CTTGAGCAGC TCTGAGGAGG AGCTGGACCT 720
 CAGATGCCCT TCCCAGGCTT TAGAGGAGCC TGAAGATGCA GAAGTCTTCA CAGAATCAAG 780
 CAGTTATGTT GAAAAGTACA ACACCTTCTGA TGATTGCAGC AGCTCAGAGG AAGACCTGCC 840
 TCTCAGACAC CTTGCTCAGG CCTTGGGAAA GCCCAAAAAC CAACAAGAAG TCTCTCTGTC 900
 TTCAATAAAT ACTCCTGAAG AGCAGAATGA TTTTATGCAG CAGCTGCCTT CCAGATGCC 960
 70 TTCTCAGCCC ATTTATGAAT CTACTGTTCA GCAACAAGTC CCCACCACTT CAGTGGGCAC 1020
 TTCTATAAAA CAGAGCGATT CCGTGGAGCC AATCCCTCCA AGACACCTT TCCAGCCATG 1080
 GGTGAACCCCT AAAGTGGAGC AAGAAGTTTC CTCATCTCCA AAGAGCATGG CTGTTGAAGA 1140
 GAGCAATTTCT ATGAAGCCTC TGCCCTCTAA ACTTCTTTC CAGCCCTTGA TGAATCCTAA 1200
 AGTTCAACAA AACATGTTCT CAGGTTTCAGA GGACATTGCT GTTGAGAGAG TCATTCTGT 1260
 75 GGAGCCACTA CTCCCAGAT ATTTCTCTCA GTCTTGACA GATCCTCAA TCCGGCAAAT 1320
 CTCAGAAAAG ACAGCTGTTG AGGAAGGCAC TTATGTGGAA CCGCTGCCTC CCAGATGCC 1380
 TTCCCAGCCC TCGGAGAGGC CTAAGTTCTT GGAATCAATG AGTACTTCTG CAGAATGGAG 1440
 CAGTCTGTG GCACCAACAC CTTCCAAATA CACTTCCCG CCAATGGTGA CCCATAAAT 1500
 80 TGAGGAACCTG TATCAACTCT CTGCACATCC AGAAAGCACT ACTGTTGAAG AGGACATTT 1560
 TAAGGAGCAG CTGCTTCCA GACATCTTTC CCAGTTGACT GTGGGAAATA AAGTCCAGCA 1620
 ACTGTCTCA AATTTCGAGC GGGCTGCTAT TGAGGCAGAC ATTTCTGGGA GTCCATTGCC 1680

5 TCCCAATAT GCTACCCAGT TCTTAAAGAG GTCTAAAGTT CAGGAAATGA CCTCAGCACT 1740
 AGAGAAAATG GCTGTGTAGAG GCACATCTTAA CAAATCACCAG ATTCACAGGC GTCCGACCCA 1800
 GTCATTCGTG AAATTTATGG CACAGCAAAAT CTTTTCAGAG AGCTCTGCTC TTAAGAGGGG 1860
 CAGTGATGTG GCACCTCTGC CTCCCAATCT TCCTTCCAAA TCTTTATCAA AGCCTGAAGT 1920
 CAAGCACCAA GTTTTCTCAG ATTACAGGAG TGCTAATCCT AAGGGAGGCA TTTCTTCAAA 1980
 GATGCTACCT ATGAAGCACC CTTTACAGTC CTTGGGGAGG CCTGAAGACC CACAGAAAGT 2040
 TTTCTCTTAT TCAGAGAGAG CTCTCGGGAA GTGCAGCAGT TTTAAAGAGC AGCTGTCTCC 2100
 CAGGACAGCT TCCCAAGCCT TGAGGAAACC TGAGTATGAG CAAAAAGTCT CCCCTGTTTC 2160
 10 TGCCAGTTCT CCTAAGAGT GGAGGAATTC TAAAAAGCAG CTGCCTCCCA AACATTCTTC 2220
 CCAAGCCTCA GATAGGTCTA AATTCCAGCC ACAGATGTCA TCAAAGGGCC CAGTGAATGT 2280
 ACCTGTAAAG CAGAGCAGCG GTGAGAAACA CCTGCCATCA AGTAGTCTCT TCCAGCAACA 2340
 GGTTTATTCA AGTTCTGTGA ATGCTGCTGC TAGGCGATCT GTTTTGTAGA GCAATTCTGA 2400
 CAATTTGGTTC CTAGGAAGAG ATGAAGCTTT TGCAATCAAA ACCAAGAAAT TCAGCCAAGG 2460
 15 TTCCAAAAAC CCCATAAAGA GCATTCACAG CCCTGCTACC AAACCTGGGA AGTTCACCAT 2520
 TGCTCCTGTC AGGCAAAACAT CCACCTCTGG GGGCATTTCAC TCTAAGAAAG AAGATCTTGA 2580
 GAGTGGTATG GGTAAATAATA ACCAGCATGC AAACCTATCC AATCAGGATG ATGTTGAAAA 2640
 GCTTTTGTGA GTTCGACTGA AAAGAGCCCC TCCTTCGCAG AAGTATAAGA GTGAGAAACA 2700
 AGATAACTTC ACCCAGCTTG CTTTCAGTGCC CTCGGGCCCA ATTTCAATCCT CTGTAGGCAG 2760
 20 GGGACATAAA ATCAGAAAGA CTTTCCAGGG GCTCCTGGAT GCTGCAGGGA ACCTCACCAA 2820
 AATACTTTAC GTTCAGATA AGCAACAGAG CAGGCCCAAA TCTGAAAGCA TGGCCAGAA 2880
 GCAACCTGCT TGCAAGACCC CAGGAAAGCC TGCTGGTCAA CAGTCAGATT ATGCTGTCTC 2940
 AGAGCCGGTT TGGATAACTA TGGCAAAGCA GAAGCAGAAG AGTTTCAAGG CCCACATTTT 3000
 TGTGAAAGAG CTGAAAACTA AGAGCAATGC TGGAGCCGAT GCTGAGACTA AGGAGCCTAA 3060
 25 ATATGAGGGA GTGGCTCTG CAAATGAAAA CCAACCTAAA AAGATGTTC CTTCCAGTGT 3120
 CCATAAACAG GAGAAGACAG CACAGATGAA GCCACCTAAG CCTACAAAA CAGTTGGATT 3180
 TGAAGCTCAG AAGTACTGCG AAGTTCTGCG CATGGA AAAA GAAACCAAAC GATCTTCAAC 3240
 TCTCCACGCC AAGTTCCAGA ACCCAGTTGA GCCAATTGAG CCTGTCTGGT TCTCACTGGC 3300
 CAGGAAGAAA GCCAAAGCAT GGAGCCACAT GGCAGAAATC ACGCAATAAA GAGCTCTTGT 3360
 30 GTGAGCATC AGCATTTATT TTAATTAGTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 3420
 CTCCGCTGCT TACCAGATT GGAGTGCAGT GGCCTGATCT CCGCTCACTG CAAGCTCCGC 3480
 TCCCGGGTT CACGCCACTC TCCCGCTTCA GTCTCCGAC TAGCTGGGAC TACAGGCGCC 3540
 CGCCATCAG CCCGCTAAT TTTGTTTCG TATTTTAGT AGAGACGGGG TTTCAACATG 3600
 TTTGGCCAGA TGGTCTTAT CTCTGACCT CGTGATCCGC CCGCTCAGC CTCCCAAGG 3660
 35 CTGGGATTAC AGCGGTGAGC CACCGCGCCC GGCCAAGCAT CAGCGTTTAA AATGATAATT 3720
 GCTAATAGCT GTATTAAATC TATGTAGTGA TCTTTTACT GTGACCATT GTATTAAAGCA 3780
 AAATAAGTAT TAAGCAAACT AAGAATTTAT TAAGCAAAAT AAGAATTTAT TAAGCAAAAT 3840
 AGCCTTAGAA ATGCAAAATTA AAACATAATT ATTTGAAATG AATAAATGCC ATGAATGCTT 3900
 AACCTTCCAC GTAGTCACTG CCAGCACCCA GAAACCCAGC ATTTCTCTTA TTAACCTAT 3960
 40 CGAAACATT TGCACCTGCTG TAAATTTGCA AAATCTTTAA CTTTGGACAA TGTGCTTTAG 4020
 AAGGAGAAA GCCTAACCAT TTTGTTGGAG CAACATAGAAA ATTTGATCTT CCCTCAACCA 4080
 AATAAAGTAA TTCTAATGGA AACATTCAGA TGATTTGACC TAAAGATTGG CCTTTAGGTT 4140
 TTAATGAGCT AGATAGATGC CGCAATTATT TGGTTGTTGC TCTAAGCTTT GCAAGGGATC 4200
 CTAAAGAGG CGGTGGAAGT GAAATTTCTG GGTCTCCAG AAAATTTCTG CACAGCCAGT 4260
 45 TCTCCAATCA GCCTATACCC CCTTGAAACA TCTTCCCTGT GTCCCTGGGG GCCCCTGATG 4320
 CTTTCTCCTT GGGTGATAGT AACATGCAGA GCACCTACAC AAGCTCTCCT CTTTGGACAT 4380
 ACCCCAGCTG GACCTGTGAC AGGCTTGCTG GTAGCGAGCA CCTCCCTATG ACGCAGAAATG 4440
 CTTCTTGGGA ATTAATCTAC TCTCTGGAG GGTAGTCCA TCAATGTTT GCTTCTTGTG 4500
 CCAATACTAC TGTGACCTC TCTGATCGCA CAGAAATCAC TGCCTATCAC ATATATCTCTG 4560
 50 TTAAGCAGTG AAGACCTTAT TGAATTTAGA GTTCTACAGA TGCCAAAAGC TGTACTTTCC 4620
 ATCAGGCGA TGGCAAGCTT ACTGCTTGA TGCACATCTG GAGCCACTGG AGCTCCTTCC 4680
 TCTCTGGTTC CAGCATTAAG GTGAGAACT CCATGTAGCT TCTTGTCTCT TCCCTCAGC 4740
 TGTCTTTGCT TCACAAGGTT TTAGCCAAA GCAAGAGTGC AATCCCAAAG CCACAGAGAA 4800
 ATGAACCTTC CGCTACCTGG AAGCTTTAAG TGAGTAAATC AGCTTTTCCC CTCTCATTTCC 4860
 55 TAGAGGCACA CACCTCAAAA GTTACTAGGC TGGAGAGACC CTACCTTTCCA GTGACCCACT 4920
 CATCCCCAG CCACGGAGAA GAGGGAAGAC CAAAAAGGGA GAGTGAGAAA GAGGATGAGA 4980
 GGGATGGTCA GCTGTGAGGG GAGGGGGCAA GTGGCCAGC AAATGTTGAT GCCTCCCTTC 5040
 CCATCTGCC ACACGGTCTT TTTCTTTTGT AGCAGAGCT CCATTAATAA CTCTCGGCT 5100
 GAGGATGAAG ATGTAGGCAC CTTTACCCCC AGAGCCAGTT CCTTAATTGG CTGGCTTTCT 5160
 60 GAGATGCAGA CCACCTTAGA ATCTCATCTA GGTTCACATG AAGTTAGTTA AATCTTCTCT 5220
 TCTCTGTCTT TCTCTCTATT CCATCCCCCA AACCCACCAA ACCTAAGGG AGAGCTCCCT 5280
 TGGATGTCT GGGCAGTAAA CCTAGCTCAT TTTTCTAGGA GACCCAGAAG TGACTTCTGA 5340
 GTAGTTATCA CTGTCTCTGC CTCTGTTTACA CTGTGCTGCT TTGCTTAAAC AGAAATGCAG 5400
 65 GCTTGGACAT CTGACTGTGC CTTTATATTC TGAGTGGGGT GCTGCCCAT GCAAAAAAAT 5460
 CCAGAGAGGT AGTGAGGTGT CAGAGCTAAA CACTTGGTGC TGGGTTTGTG TGATGCTGGT 5520
 ATAATGTGAC ACAGTACAT TACATGCTAA ATTTTGCATT TTTCTATAT AACATCTATT 5580
 TTTCTGATA CTGTGCCCTTT GCCATTTTGA TAATGCTATT TTGATTGAGT GAATTTTATT 5640
 TCTTTGTAT TCCCATAGTG AACAAATAT TAAGGTAGAT GCGCTTTATC TGGGTACTCC 5700
 70 TGGTAGATTA GCTGTATGAC CTCCCTTCCC TTTTTCACAG TGAACCTGTA TTCAGTTATT 5760
 GTCACCTGTA GAACCTCTCA ATAACAATTT CTTTTCACAG GTTAACAACA CAGCTGTTAC 5820
 ACCCTCCCTC CTTTTTTCCA CAGTGAACTT GTATTCAGCT ATTTCTCACTC TGAGAACTCT 5880
 CCAATAACAA TTTCTTTTCC ACAGTTAACA ACAAGTTCT GTTTTAAAT GAAGAGATTA 5940
 AGTTCTTTT AAATGCCATA AGGCATATTC TGACAACCTT TCTACTTCTT TAACCTTTT 6000
 GATTTAAGAT ATATGCAAG CAAATAAAT CAATAAAGCC T

SEQ ID NO:48 PDG5 Protein sequence

Protein Accession #: BAA86524

1 11 21 31 41 51
 EQPTTSQPET TTPQGLLSDK DMGRRNAGI DFGSRKASAA QPIPENMDNS MVSDFQPYHE 60

DAASGAEKTE ARASLSLMVE SLSTTQEEAI LSVAAEAQVF MNPSHIQLED QEAFSFDLQK 120
 AQSKMESAQD VQTICKEKPS GNVHQTFTAS VLGMTSTTAK GDVYAKTLPP RSLFQSSSRKP 180
 DAEVSSDSE NIPEEGDGSE ELAHGHSSQS LGKFEDEQEV FSEKSFVED LSSSEEBDL 240
 RCLSQALEEP EDAEVFTSS SYVEKYNTSD DCSSEEDLP LRHPAQALGK PKNQEVSSA 300
 SNNTPEEQND FMQQLPSRCP SQPIMNPTVQ QQVPTSSVGT SIQSDSVEP IPPRHFPQFW 360
 VNPKEQEVSS SSPKSMAVEE SISMKPLPPK LLCQPLMNPV VQNMFSGSE DIAVERVISV 420
 EPLLPRYSFQ SLTDPQIRQI SESTAVEEGT YVEPLPPRCL SQPSERPKFL DSMSTSAEWS 480
 SPVAPTPSKY TSPFWPTPKF EELYQLSAHP ESTTVEEDIS KEQLLPRHLS QLTVGKQVQ 540
 LSSNFERAAI EADISGSLP PQYATQFLKR SKVQEMTSRL EKMAVEGTSN KSPIPRRPTQ 600
 SFVKFMAQQI FSESSALKRG SDVAPLPENL PSKSLSKPEV KHQVFSDSGS ANPKGGISSK 660
 MLPKHPQLQS LGRPEDPQKV FSYSERAPGK CSSFKEQLSP RQLSQALRKP EYEQKVSPVS 720
 ASSPKWEWNS KKQLPPKHSS QASDRSKFPQ QMSSKGFVNV FVKQSSGEKH LPSSSPFQQQ 780
 VHSSSVNAAA RRSVFESNSD NWFLGRDEAF AIKTKKFSQG SKNPIKSIPA PATKPGKFTI 840
 APVRQTSTSG GIYSKKEDLE SGDGNNNQHA NLSNQDDVEK LFGVRLKRAP PSQKYKSEKQ 900
 DNFTQLASVP SGPISSSVGR GHKIRSTSQG LLDAAAGNLTK ISYVADKQOS RPKSESMACK 960
 QPACKTPGPK AGQQSDYAVS EPVWITMAKQ KQKSFKAHIS VKELKTKSNA GADAETKEPK 1020
 YEGAGSANEN QPKMFTSSV HKQEKTAQMK PPKPTKSVGF EAQKILQVPA MEKETKRST 1080
 LPAKQNVFE FIEPVWFLSA RKKAKAWSHM AEITQ

SEQ ID NO:49 PAB7 DNA SEQUENCE

Nucleic Acid Accession #: D87742

Coding sequence: 208-3582 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 GCTTTCCTTT CTAAAGTAGA AGAGGATGAT TATCCCTCTG AAGAACTACT AGAGGATGAA 60
 AACGCTATAA ATGCAAAACG GTCTAAAGAA AAAAACCTCG GGAATCAGGG CAGGCAGTTT 120
 GATGTTAATC TGCAAGTCCC TGACAGAGCA GTTTTAGGGA CCATTTCATCC AGATCCAGAA 180
 ATTGAAGAAA GCAAGCAAGA AACTAGTATG ATTTTGGATA GTGAAAAAAC AAGTGAGACT 240
 GCTGCCAAAG GGGTCAACAC AGGAGGCAGG GAACCAATA CAATGGTGGG AAAAGAACCG 300
 CCTCTGGCAG ATAAGAAAGC ACAGAGACCA TTTGAACGAA GTGACTTTTC TGACAGCATA 360
 AAAATTCAGA CTCAGAAATT AGGTGAAGTG TTTCAAGAATA AAGATTCTGA TTATCTGAAG 420
 AACGACAACC CTGAGGAACA TCAGAAGACC TCAGGGCTTG CAGGGGAGCC TGAGGGAGAA 480
 CTCTCAAAAG AGGACCATGG GAACACAGAG AAGTACATGG GCACAGAAAG CCAGGGGTCT 540
 GCTGCTGCAG AACCTGAAGA TGACTCGTTC CACTGGACTC CACATACAAG TGTAGAGCCA 600
 GGGCATAGTG ACAAGAGGGA GGACTTACTT ATCATAAGCA GCTTCTTTAA AGAACACAG 660
 TCTTTGCAGC GGTTCAGAA GTACTTTAAT GTCCATGAGC TGAAGCCTT GCTACAAGAA 720
 ATGTCATCAA AACTGAAATC AGCGCAGCAG GAGAGCCTGC CCTATAATAT GGAAGAAATC 780
 CTAGATAAGG TCTTCCGTGC TTCTGAGTCA CAAATTCGTA GCATAGCAGA AAAAATGCTT 840
 GATACTCGTG TGGCTGAAA TAGAGATCTG GGAATGAACG AAAATAACAT ATTTGAAGAG 900
 GCTGCAGTGC TTGATGACAT TCAAGACCTC ATCTATTTTG TCAGGTACAA GCACTCCACA 960
 GCAGAGGAGA CAGCCACACT GGTGATGGCA CCACCTCTAG AGGAAGCCTT GGGTGGAGCA 1020
 ATGGAAGAGA TGCACCAACT GCATGAAGAT AATTTCTCAC GAGAGAAGAC AGCAGAACTT 1080
 AATGTGCAGG TTCTGTAAGA ACCCAACCAC TTGGACCAAC GTGTGATTGG GGACACTCAT 1140
 45 GCCTCAGAAG TGTCACAGAA GCCAAATACT GAGAAAGACC TGGACCCAGG GCCAGTTACA 1200
 ACAGAAGACA CTCTATTGGA TGCTATTGAT GCAACAAGC AACCAGAGAC AGCCGCGGAA 1260
 GAGCCGGCAA GTGTCAACAC TTTGGAAAAC GCAATCCTTC TAATATATTC ATTCATGTTT 1320
 TATTTAACTA AGTCGCTAGT TGCTACATTG CCTGATGATG TTCAGCCTGG GCCTGATTTT 1380
 50 TATGGACTGC CATGGAAACC TGATTTATTC ACTGCCTTCT TGGGAATTGC TTCGTTTGCC 1440
 ATTTTCTTAT GGGAAGCTGT CCTTGTGTGT AAGGATAGAG TATATCAAGT CACGGAACAG 1500
 CAAATTTCTG AGAAGTTGAA GACTATCATG AAAGAAAATA CAGAAGTTGT ACAAATAATTG 1560
 TCAATATTAT AACAGAAGAT CAAGGAATCA AAGAAACATG TTCAGGAAAC CAGGAAACAA 1620
 AATATGATTC TCTCTGATGA AGCAATTAAA TATAAGGATA AAATCAAGAC ACTTGAAAAA 1680
 55 AATCAGGAAA TTCTGGATGA CACAGCTAAA AATCTTCGTG TTATGCTAGA ATCTGAGAGA 1740
 GAACAGAATG TCAAGATACA GGACTTGATA TCAGAAAACA AGAAATCTAT AGAGAAGTTA 1800
 AAGGATGTTA TTTCAATGAA TGCCCTCAGAA TTTTCAGAGG TTCAGATTGC ACTTAATGAA 1860
 GCTAAGCTTA GTGAAGAGAA GGTGAAGTCT GAATGCCATC GGGTTCAAGA AGAAATGCT 1920
 AGGCTTAAGA AGAAAAAGA GCAGTTGCAG CAGGAAATCG AAGACTGGAG TAAATTACAT 1980
 60 GCTGAGCTCA GTGAGCAAAT CAAATCATTT GAGAAGTCTC AGAAAGATTT GGAAGTAGCT 2040
 CTTACTCACA AGGATGATAA TATTAATGCT TTGACTAACT GCATTACACA GTTGAATCTG 2100
 TTAGAGTGTG AATCTGAATC TGAGGGTCAA AATAAAGGTG GAAATGATTG AGATGAATTA 2160
 GCAANTGGAG AAGTGGGAGG TGACCGGAAT GAGAAGATGA AAAATCAAAAT TAAGCAGATG 2220
 ATGGATGTCT CTCGGACACA GACTGCAATA TCGGTAGTTG AAGAGGATCT AAAGCTTTTA 2280
 65 CAGCTTAAAG TAAGAGCCTC CGTGTCCACT AAATGTAACC TGGAAAGACCA GGTAAGAGAA 2340
 TTGGAAGATG ACCGCAACTC ACTACAAGCT GCCAAAGCTG GACTGGAAGA TGAATGCAAA 2400
 ACCTTGAGGC AGAAAGTTGA GATTCTGAAT GAGCTCTATC AGCAGAAGGA GATGGCTTTG 2460
 CAAAAGAAAC TGAGTCAAGA AGAGTATGAA CGGCAAGAAA GAGAGCAGAG GCTGTGAGCT 2520
 GCAGATGAAA AGGCAGTTTC GGTGTCAGAG GAAGTAAAAA CTACACAGCG GAGAAATTGA 2580
 70 GAAATGGAGG ATGAATTACA GAAGACAGAG CGGTCAATTA AAAACCAGAT CGTACCCAT 2640
 GAGAAAGAAG CTCATGAAAA CTGGCTCAAA GCTCGTGCTG CAGAAAGAGC TATAGCTGAA 2700
 GAGAAAAGGG AAGCTGCCAA TTTGAGACAC AAATTATTAG AATTAAACACA AAAGATGGCA 2760
 ATGCTGCAAG AAGAACCTGT GATTGTAAAA CCAATGCCAG GAAAAACAAA TACACAAAAC 2820
 CCTCCACGGA GAGGTCTCTT GAGCCAGAAAT GGCTCTTTTG GCCCATCCCC TGTGAGTGGT 2880
 75 GGAGAATGCT CCCCTCAATT GACAGTGGAG CCACCCGTA GACCTCTCTC TGCTACTCTC 2940
 AATCGAAGAG ATATGCCATG AAGTGAATTT GGATCAGTGG ACGGGCCTCT ACCTCATCCT 3000
 CGATGGTCAG CTGAGGCATC TGGGAAACCC TCTCCTTCTG ATCCAGGATC TGGTACAGCT 3060
 ACCATGATGA ACAGCAGCTC AAGAGGCTCT TCCCCTACCA GGGTACTCGA TGAAGGCAAG 3120
 GTTAATATGG CTCCAAAGGG GCCCCTCTCT TCCCAGGAG TCCCTCTCAT GAGCACCCCC 3180
 80 ATTGGGAGGCC TGCTACCAAC ACCCATTCGA TATGGACCAC CACCTCAGCT CTGCGGACCT 3240
 TTTGGGCCCTC GGCACTTCTC TCCACCTTTT GGCCCTGGTA TGGTCCACC ACTAGGCTTA 3300

	AGAGAATTTG	CACCAGGCGT	TCCACCAGGA	AGACGGGACC	TGCTCTCCA	CCCTCGGGGA	3360
	TTTTTACCTG	GACACGCACC	ATTTAGACCT	TTAGGTTTAC	TGGGCCCCAG	AGAGTACTTT	3420
	ATTCCCTGGT	CCCGATTACC	ACCCCCAACC	CATGGTCCCC	AGGAATACCC	ACCACCACCT	3480
	GCTGTAAGAG	ACTTACTGCC	GTCAGGCTCT	AGAGATGAGC	CTCCACCTGC	CTCTCAGAGC	3540
5	ACTAGCCAGG	ACTGTTTACA	GGCTTTAAAA	CAGAGCCCAT	AAAACATATGA	CCTCTGAGGT	3600
	TTCATTTGGA	AGAAAGTGTA	CTGTGCATTA	TCCATTACAG	TAAAGGATTT	CATTGGCTTC	3660
	AAAATCCAAA	AGTTTATTTT	AAAAGGTTTG	TGTTTAGAAC	TAAAGTGCCT	TGGCAGTGTG	3720
	CATTTTGTAG	CCAAACAATT	CAAAAATGTC	ATTTCTTCCC	TAAATAAAAA	TCACCTTTTA	3780
	AGCTAGAGCG	TCCTTACAAC	TTTGAAATGT	GCAATAAAGA	ATACCTGTGT	TTTAGCTAAT	3840
10	GTAGCATATG	TAATTGCAAA	ATGATTTAGA	ATGTCATGAA	AAATATGAAC	ATTTCTCTGT	3900
	GAAATGCTTT	AAGAACATGT	ATTTCCATTA	TCCTATTTTT	AGTGACACCC	AGCTGAATAC	3960
	GGAGCAATGG	TGTTTATAAG	CGTTTTTTTA	AACTATCTGG	TCACAAAGAC	TGTTACGCTA	4020
	AAAATGTTTA	CTAAAAGATC	ACTAAACTAT	CTCCCTCTCT	GCTGAAGTTC	TTTGTAGTAA	4080
	TAGCTCATAA	AAATTTGTTT	ATTAATATTT	CCCAAGTGTG	TGTTGACTCA	TTGGACTGTT	4140
15	ATGAGGCTTG	TGCCATTTGG	GGAACATGTA	AACTCAGGCT	CCCAGAACTG	AAGATGGTGG	4200
	CTGGTGGCAC	ACTTCCGGGT	GCTCCTCCGT	CACCTGTGAA	CTCTACAAGT	GATGTCTTTT	4260
	TATTTCAAAG	AAGTTTATTT	CCCACTTGTA	TAGCATTCAC	ATGCTTTCTT	TACGATCCTC	4320
	ATTGTCTATT	TGAGAATGGT	TTTCTGAGAG	TGAGTTTACA	TTAGTAGCAA	GAGTTGTTTG	4380
	ACCTGATGTT	CCATTGTTTT	TACCATTCTT	GTAGAAAAAG	GGTGACACAC	AGAAAAATGA	4440
20	AAATGATGTG	TCATGGCCAT	AAAAGTATAG	AAATCTTTAA	AAATTTTAAA	ATGTACAGTC	4500
	CCTTATCTAT	CTTTCCCATT	CCTTGCCACT	GATTTTGTAG	GAATATAATA	AAAAGATTGG	4560
	AAAGTATATA	TGCCATGAGA	AAGAATGATT	TAGGACTGTG	AGGGTTATAA	CATGCCCTAG	4620
	GTCAGCAACC	AAGGGTTGAA	ATCAGTTCTG	TTTTAGGGGG	AAATGGGGGG	GGCGACAGAT	4680
25	ATATTTCCAA	AATTAATATT	AATTAATATT	TAAACGTTGG	TGTTTTTATT	TAAAAATCAG	4740
	TAACTAACCA	TCGTGAATTG	CACCATACTT	AAAGTCTTAT	CCATTACTAC	ACTGTCTTTA	4800
	AAACAATGTT	CTTTTAAATA	CTCTACAACG	TTTCTAAGAA	CGAAGTTCAG	ACATTTTAAAT	4860
	TACAGTAATA	ATAGCACTCC	TTTTAAGGAG	TTTCAGATCC	ACACTAAAAC	TAAAAATCATA	4920
	AAAGGCTGAT	ACTTTTGTGT	GCTGCTAGGC	TATATCTTTC	CATTCTTTGA	AGTCCTATGA	4980
30	TGTAATATTT	TTGAAACCTA	GTGTATGTCT	TGCTACTGTT	GTGATATTTA	ATCGATTAAAG	5040
	AAATACCTTG	AAAAAGGAGC	AAAAGCTTCA	ATGTGAAACA	ATTTCTCTCT	TTTATACTAA	5100
	ACAACGAAG	ATAGATAGTT	TAGAAAGATA	AGGACCTTTG	AAAGAAGACA	ACTCTGTCAA	5160
	AGTTCATAAG	GAATATAAAA	ATTCCTCAGG	AAAAGAGAAT	TCAATCTATA	TGTCCTCCCG	5220
	TTTAATATCA	AGAATAGAAG	AAATTAAGAG	GAAAACCTCA	CAGAAGAGCA	TAGGCCACTT	5280
35	TTAGCCATGT	AAAAATAAGA	TTAAGTCACA	AATACAACCT	TTGAATTTAC	CTGTCAATAT	5340
	CTCTTTAGGA	CACAAAACAA	TGCTGAAGTT	AATATAATTT	CTAATTTTAA	ATGTCATTTA	5400
	AGTGTAGATT	TGCCATCTTA	GGAAGGTAAG	TAGGAAAGGT	AAATTAATC	TATTTTAAAA	5460
	ATTCAAAAAA	TTAGAGTATT	TTTCCCTCT	AAAGCCTTTT	TTGGTGATTA	TTCTGTATCT	5520
	GACATAATTG	AGAAACTGGT	AAGCTGTAAA	GATTCCAGTG	TAGCTTCTCT	GAGAAGTTGT	5580
40	GAGCCAGTCC	ATAACTGCCT	CCTCACATCC	ATCTGATTGC	ACCATTTCTG	CAGCAAAACC	5640
	CAAAGCAGGG	TGCCAATATG	CAGATGGCAT	AGGGAGTATC	ATCCCTCAGC	CAAAATCACTT	5700
	TTCCATCTCT	AAAGTTTGA	CTATTTTGA	AGTCATCTCC	AACTAATTGT	GTCTGGATT	5760
	AGTTGTCTAA	ATTGTCTTAT	TTATTTATGA	AGCAGCAATA	TTAGCCCTGA	AAGCATTTCT	5820
45	GCCATAGTTG	TTGTAGTTAT	ATCGCCAATG	GCTGATTTT	TTTATTTGAA	AGTAAATTTA	5880
	AGTAATTCGT	GGGATGTGGT	ATATTCTGTG	TCAACTTCAA	GATAATCACT	CATTTTCTCG	5940
	TTATATTTCAG	GTCTGAATTA	AAGTTAAGTT	AATCAC			

SEQ ID NO:50 PAB7 Protein sequence

Protein Accession #: BAA13448

50	1	11	21	31	41	51	
	AFLSKVEEDD	YPSEELLEDE	NAINAKRSKE	KNPGNQGRQF	DVNLQVPDRA	VLGTIHPDPE	60
	IEESKQETSM	ILDSEKTSET	AAKGVNTGGR	EPNTMVEKER	PLADKKAQRP	FERSDFSISI	120
55	KIQTPELGEV	FQNKSDYLK	NDNPEEHLKT	SLGAGEPEGE	LSKEDHGNTE	KYMGTESQGS	180
	AAAEPEDDSF	HWTPTSVEP	GHSDKREDLL	IISPFKEQQ	SLQRFPQKYFN	VHELEALLQE	240
	MSSKLKSAQQ	ESLPYNMEKV	LDKVPFRASE	QILSLAEKML	DTRVAENRDL	GMNENNIFEE	300
	AAVLDDIIDL	IYFVRYKHST	AEETATLVMA	PPLBGLGGA	MEEMQPLHED	NFSREKTAEI	360
	NVQVPEEPFH	LDQRVIGDTH	ASEVSKPKNT	EKDLDFGPVT	TEDTFMDAID	ANKQPETAEE	420
60	EPASVTPLEN	AILLIYSFMF	YLTKSLVATL	PDDVQPGPDF	YGLPWKPVFI	TAFLGIASFA	480
	IFLWRTVLVV	KDRVYQVTEQ	QISEKLKTIM	KENTELVQKL	SNYEQKIKES	KKHVQETRRQ	540
	NMILSDEAIK	YKDKIKTLEK	NQEILDDTAK	NLRVMLESER	EQNVKNQDLI	SENKKSIEKL	600
	KDVISMNASE	FSEVQIALNE	AKLSEEVKVS	ECHRVQEENA	RLKKKKEQLQ	QEIEDWSKLH	660
65	AELSEQIKSF	EKSQKDLEVA	LTHKDDNINA	LJNCITQLNL	LECESESEGO	NKGMDNDEL	720
	ANGEVGGDRN	EKMKNQIKQM	MDVSRQTQAI	SVVEEDLKLL	QLKLRSVST	KCNLEDQVKK	780
	LEDDRNLSQA	AKAGLEDECK	TLRQKVEILN	ELYQQKEMAL	QKKLSQEEYE	RQEREHLISA	840
	ADEKAVSAAE	EVKTYKRRIE	EMEDLQKTE	RSFKNQIATH	EKKAHENWLK	ARAAERATAE	900
	EKREANLRH	KLLELTQKMA	MLQEEPVIK	PMPGKENTQN	PFRRGPLSQN	GSFGSPSPVSG	960
70	GECSPPLTVE	PPVRPLSATL	NRRDMPRSEF	GSVDGPLPHF	RWSAEASGKP	SPSPDPSGTA	1020
	TMMNSSSRGS	SPTRVLDEGK	VNMAPKGGPP	FPGVPLMSTP	MGGVPVPPPIR	YGPPLQLCGP	1080
	FGPRLPPPPF	GGGMRPFLGL	REFAPGVPPG	RRDLPLHPRG	FLPGHAPFRP	LGSLGPREYF	1140
	IPGTRLPPPT	HGFQEYPPPP	AVRDLPLPSGS	RDEPPFASQS	TSQDCSQALK	QSP	

SEQ ID NO:51 PAB9 DNA SEQUENCE

Nucleic Acid Accession #: NM_006457

Coding sequence: 84-1874 (underlined sequences correspond to start and stop codons)

80	1	11	21	31	41	51	
	AGACTGAGGC	GGAGGCAGCC	CCGCGCCGCG	CCGGACCGGA	GCATATTTC	TTTCTGTCA	60

5 TTTGACTTTG AGCCATTAGA ACCATGAGCA ACTACAGTGT GTCACTGGTT GGCCAGCTC 120
 CTGGGGTTT CCGGCTGCAG GCGGTAAGG ATTTCAACAT GCCTCTGACA ATCTCTAGTC 180
 TAAAAGATGG CGGCAAGGCA GCCCAGGCAA ATGTAAGAAT AGGCGATGTG GTTCTCAGCA 240
 TTGATGGAAT AAATGCACAA GGAATGACTC ATCTTGAAGC CCAGAATAAG ATTAAGGGTT 300
 GTACAGGCTC TTTGAATATG ACTCTGCAAA GAGCATCTGC TGCACCCAAG CCTGAGCCGG 360
 TTTCTGTTC AAGGGGAGAA CCTAAAGAAG TAGTTAAACC TGTGCCCATT ACATCTCCTG 420
 CTGTGTCCAA AGTCTACTCC ACAACAACA TGGCCTACAA TAAGGCACCA CGGCCTTTTG 480
 GTTCTGTGTC TTCACCAAAA GTCACATCCA TCCCATCACC ATCTGTCTGCC TTCACCCAG 540
 10 CCCATGCGAC CACCTCATCA CATGCTTCCC CTTACCCCGT GGCTGCCGTC ACTCCTCCCC 600
 TGTTCGCTGC ATCTGGACTG CATGCTAATG CCAATCTTAG TGCTGACCAG TCTCCATCTG 660
 CACTGAGCGC TGGTAAACT GCAGTTAATG TCCCACGGCA GCCCAGAGTC ACCAGCGTGT 720
 GTTCCGAGAC TTCTCAGGAG CTAGCAGAGG GACAGAGAAG AGGATCCAG GGTGACAGTA 780
 AACAGCAAAA TGGCCCAACA AGAAAACACA TTGTGGAGCG CTATACAGAG TTTTATCATG 840
 15 TACCCACTCA CAGTGAATGCC AGCAAGAAGA GACTGATTGA GGATACTGAA GACTGGCGTC 900
 CAAGAACTGG AACAACTCAG TCTCGCTCTT TCCGAATCCT TGCCACAGATC ACTGGGACTG 960
 AACATTGTA AGAATCTGAA GCCGATAATA CAAAGAAGGC AAATAACTCT CAGGAGCCCT 1020
 CTCCGCAAGT GGCTTCCCTG GTAGCTTCCA CACGAGCAT GCCCGAGAGC CTGGACAGCC 1080
 CAACCTCTGG CAGACCAGGG GTTACCAGCC TCACAACATGC AGCTGCCTTC AAGCCTGTAG 1140
 GATCCACTGG CGTCATCAAG TCACCAAGCT GGCAACGGCC AAACCAAGGA GTACCTTCCA 1200
 20 CTGGAAGAAT CTCAACACGC GCTACTTACT CAGGATCAGT GGCACACGCC AACTCAGCTT 1260
 TGGGACAAAC CCAGCAAGT GACCAGGACA CTTTAGTGCA AAGAGCTGAG CACATTCCAG 1320
 CAGGGAACGC AACTCCGATG TCGGCCCATT GTAACCAAGT CATCAGAGGA CCATTCTTAG 1380
 TGGCACTGGG GAAATCTTGG CACCCAGAAG AATTCAACTG CGCTCACTGC AAAAATACAA 1440
 25 TGGCCTACAT TGGATTTGTA GAGGAGAAAG GAGCCCTGTA TTGTGAGCTG TGCTATGAGA 1500
 AATTCTTTGC CCTGAATGT GGTGAGTCCC AAAGGAAGAT CCTTGGAGAA GTCATCAATG 1560
 CGTTGAAACA AACTTGGCAT GTTTCCTGTT TTGTGTGTGT AGCCTGTGGA AAGCCCATTC 1620
 GGAACAATGT TTTTCACTTG GAGGATGGTG AACCTACTG TGAGACTGAT TATPATGCC 1680
 TCTTTGGTAC TATATGCCAT GGATGTGAAT TTCCCATAGA AGCTGGTGAC ATGTTCCTGG 1740
 30 AAGCTCTGGG CTACACCTGG CATGACACTT GCTTTGTATG CTCAGTGTGT TGTGAAAGTT 1800
 TGGAGGTCAC GACCTTTTTC TCCAAGAAGG ACAAGCCCTT GTGTAAGAAA CATGCTCAT 1860
 CTGTGAATTT TTGAAGTCA ACAGTTCAGG AGAAGAGAAG GAATTTGAAG AGAAAAAGGA 1920
 AAATTAATAA TACTAATTA TTTTITAGATT CAATATTTAT ATGGAGTTT GAAAAATAA 1980
 AGTGGCCCTG AAGGAATAAA TTCCAGCTTT AAAAACCAAG TCTGAGGAAA TATTTGGCTT 2040
 35 CATAAAGTAA AGAGACGGTT TGGCATTAT TATTACTTTT TCCTGTATTT TATGCCCAT 2100
 AAATAAGCTT TATAAAGACC AATTTCCTGA TGGACTATTA AATTCATCTT AGAATAAAT 2160
 AGTGAAGAAT TTAATTTTAG AATAAATAAT CCAATCTGAA ATAATTATAC CTTCCTTCT 2220
 TGTAGGTAG TTATGAGTAA ATCTGCAAAA GGCAATGAAA ATGCCCTAAA TTTTATCAAT 2280
 40 AACAGAATTA TTGTATTTAA AAAAAAATA ATACTTATCT TTAAATAGT AAATAGGATT 2340
 TTAACAGAG AATTTATCA GTAAAGGTG TCAGTTTATA AAAAATTGCT TGTAGGCTGA 2400
 GCGCGGTGGC TCACGCCTGT AATCCCAGCA CTTTGGGAGG CCAAGGTGGG TGGACCACAT 2460
 GAGGTCAGGA GTTTGAGATC AGCCTGGCCA ACATGGTGAA ACCCCATCTC TACTAAAAAT 2520
 ACAAAAATTA GCCCGACGCA GTGGCACGCG CCTGTAATCC CAGCTACTCA AGAGGCTGAG 2580
 GCACGAGAA CTCTGAACC CGGAGGGGAG AGGTGTCAGT GAGCCAAGAT CGTACCACTG 2640
 45 CACTCCAGCC TGGGTGACAG AGTGAGACTC CGTCTCCAAA AAAAACTTT GCTTGTATAT 2700
 TATTTTGGC TTACAGTGGG TCATTTCTAGT AGGAAAGGAC AATAAGATT TTTATCAAAA 2760
 TGTGTATGC CAGTAAGAGA TGTATATTC TTTTCTTATT TCTTCCCCAC CCAAAAATAA 2820
 GCTACCATAT AGCTTATAAG TCTCAAATTT TTGCCTTTTA CTAAATGTG ATTGTTCTG 2880
 50 TTCAATGTGT ATGCTTCATC ACCTATATTA GGCAAAATCC ATTTTITCCC TTGCGCTAAG 2940
 GTAAAGATT AATTAATAA TTTTGGCCTC TCATAGTTT CTCTCTCTTT AAAGAGAATA 3000
 AATAGAGGCG CAGGTGTGGT GGCTCACGCC TGTGATCCCA GCACCTTGGG AGGCCAAGAC 3060
 GGGCGGATCA TGAGGTCAAG AGATCAAGAT CATCTGGCC AACATGGTGA AACCCGTCT 3120
 CTACTAAAA TACAAAAATG AGCTGGGCAT GGTGGGGCGT GCCTGTAGTC CCATGTACTT 3180
 GGGAGGCTGA GGCAGGAAAA TTCTTGAACC CAGGAGACGG AAGTTGACAGT GAGCTGAGAT 3240
 55 CACACCCTG CACTCCAGCC TGGTGACAGA GCAAGACTCC GGCTCTT

SEQ ID NO:52 PAB9 Protein sequence

Protein Accession #: NP_006448

60 1 MSNYSVSLVG PAPWGFRLQG GKDFNMPLTI SSLKDGGKAA QANVRIGDVG LSIDGINAQG 60
 61 MTHLEAQNKI KGCTGSLNMT LQRASAAPKP EPVPVQKGEF KEVVKVPVIT SPAVSKVTST 120
 121 NMMAYNKAPR PFGSVSSPKV TSIPSPSSAF TPAHATSSH ASPSPVAAVT PFLFAASGLH 180
 181 ANANLSADQS PSALSAGKTA VNVPRQPTVT SVCSETSQEL AEGQRRGSQG DSKQONGPFR 240
 241 KHIVERYTEF YHVFTHSDAS KRRLIEDTED WRPRGTGTQS RSFRILAQIT GTEHLKESEA 300
 301 DNTKKANNQS EPSFQLASLV ASTRSMPESL DSPTSGRPGV TSLTAAAFK PVGSTGVVKS 360
 361 PSWRPNQGV PSTGRINSNA TYSGSVAPAN SALGQTQPSD QDTLVQRAEH IPAGKRTPMC 420
 421 AHCNQVIRGP FLVALGKSWH PEEFNCAHCK NIMAYIGFVE EKGALYCELC YEKFFAPECG 480
 481 RCQRKILGDM INALKQTHV SCFVCVACGK PIRNNVFHLE DGEFVCETDY YALFGTICHG 540
 541 CEFPIEAGDM FLEALGYTWH DTCFVCSVCC ESLEGQTFSS KKDKPLCKKH AHSVNF

SEQ ID NO:53 PBH7 DNA SEQUENCE

Nucleic Acid Accession #: AA431407

Coding sequence: 1-864 (underlined sequences correspond to start and stop codons)

75 1 ATGGCCAACT GTAAAAATGAC CAAAAGCATC AGGTTCCTCG CCCTGGAGCA CTGCTATACT 60
 80 GGCGGGGAGG TCGTGTGGCC CAAGGATCAG GAGGAGTGGG AAAGACGGAC GGGCCTTCTG 120
 CTCTACGAGA ACTATGGGCA GTCGGAACAG GGACTAATTT GTGCCACCTA CTGGGGAATG 180

5
10
15

```

AAGATCAAGC CGGGTTTCAT GGGGAAGGCC ACTCCACCCT ATGACGTCCA GTTTCATATG 240
GAGGCTCAGC TTGAAACTG CATTATTGTG AGCATGAACA CCGCTGACCC TGGCAGCCAG 300
GGCATCACAC ACAGCCTCTT GCTACAGGTC ATTGATGACA AGGGCAGCAT CCTGCCACCT 360
AACACAGAAG GAAACATTGG CATCAGAATC AAACCTGTCA GGCCCTGTAG CCTCTTCATG 420
TGCTATGAGG GTGACCCAGA GAAGACAGCT AAAGTGGAAT GTGGGGACTT CTACAACACT 480
GGGGACAGAG GAAAGATGGA TGAAGAGGGC TACATTGTGT TCCTGGGGAG GAGTGTATGAC 540
ATCATTAATG CCTCTGGGTA TCGCATCGGG CCTGCAGAGG TTGAAAGCGC TTTGGTGGAG 600
CACCCAGCGG TGGCGGAGTC AGCCGTGGTG GGCAGCCCAG ACCCGATTCT AGGGGAGGTG 660
GTGAAGGCTT TTATTGTCTT GACCCACAG TTCTGTCTCC ATGACAAGGA TCAGCTGACC 720
AAGGAAGTGC AGCAGCATGT CAAGTCAGTG ACAGCCCCAT ACAAGTACCC AAGGAAGGTG 780
GAGTTTGTCT CAGAGCTGCC AAAAACCATC ACTGGCAAGA TTGAACGGAA GGAAGTTCGG 840
AAAAAGGAGA CTGGTCAGAT GTAATCGGCA GTGAAGTCTG AACGCACTGC ACACCTGAGG 900
CAAATCCCTG GCCACTTTAG TCTCCCACT ATGGTGAGGA CGAGGGTGGG GCATTGAGAG 960
TGTTGATTTG GGAAGTATC AGGAGTGCCA TGATTCCAAT GTTTTCCTTC TTTTAAATTA 1020
AATTCAGTTG CTCTGCTTCC TCCAAGTCCT CTGTATCTTT AGAATTTCCT AGGTGAGCAC 1080
TCATAACGCA AGTAATAAAA TACTGATATC AACAA

```

SEQ ID NO:54 PBH7 Protein sequence

Protein Accession #: FGENESH predicted

20
25

```

1 11 21 31 41 51
| | | | |
MANCKMTKSI RFPALHCYCT GGEVVLPKDQ EEWKRRTGLL LYENYQSQSET GLICATYWGM 60
KIKPGFMGKA TFPYDVQPHM EASVENCIIIV SMNTADPGSQ GITHSLLLQV IDDKGSILPP 120
NTEGNIIGIRI KPVRFVSLFM CYEGDPEKTA KVECGDFYNT GDRGKMDEEG YICFLGRSDD 180
IINASGYRIG PAEVESALVE HPAVAESAVV GSPDPIRGEV VKAFIVLTPQ FLSHDKDQLT 240
KELQHVKSIV TAPYKYPRKV EFVSELPKTI TGKIERKELR KKETGQM

```

SEQ ID NO:55 PBJ5 DNA SEQUENCE

Nucleic Acid Accession #: AF388200

Coding sequence: 33-137 (underlined sequences correspond to start and stop codons)

35
40

```

1 11 21 31 41 51
| | | | |
GAGAGAGGGA GGCAGAAGAG GAAGTCAGAG CGATGTGCTG TGAAATCTAC TACCGTTTGC 60
TGGTTTTGAA AATGGAGAAA AAGAGTGAGG AACTGAGAAA CATGGATGGC CTTGGGAACG 120
TGGAAAAGGG TCACTGAAAT GGGACGACAT GAACTCAAGG AGGCTATTTA TGACCATGTC 180
ATTTGCAACA TGAAGAAAGC TTATCTGGAG TGAAGTAAAT TGAGACCAAC AGAGATAAGA 240
GACCCGGAGA AATCCTGGTT ACACCTGCTG AATCCTGTCA GTCCTATACT GGAGTCCTGT 300
TAATACAAAA TAATAGTAAT AATCCCCTGT TTCTTTATGT TTATGCCAAC TTCAACAAAA 360
AGAAACTTGA CTAAGAGACA ATATAAGAAC TTAATGTGTA ATTAAGAAAG AACTCTCCAC 420
CACGGGGAAT GTGAAAGGTA TATGAGTCCC TTTTCACGAT GCGATGTCAT GTCTTTTAA 480
TAAGCCATAC TTTATGTTC AATAAAGAG AATAAGCAGG A

```

SEQ ID NO:56 PBJ5 Protein sequence

Protein Accession #: AAK83352

50

```

1 11 21 31 41 51
| | | | |
MCCEIYYRLV VLKMEKKSEE LRNM DGLGNV EKGH

```

SEQ ID NO:57 PBJ7 DNA SEQUENCE

Nucleic Acid Accession #: AA876910

Coding sequence: 1-2064 (underlined sequences correspond to start and stop codons)

55
60
65
70
75
80

```

1 11 21 31 41 51
| | | | |
ATGGACAGTT GCCTGCAACA TATGAGAGAC CTACTTTACC TCCTTCAGGA GCTCAGGTGT 60
TTAAATCCAG CTACACTACT CCCTGATCCA GACTCCACTA CTCCTGTTC TACTGTCTAG 120
GATCTGTGGG AAACTACCAA AACTGGCCAA CCTGATCTTC AAGATGTGCC CCTAGAAAAG 180
GCAGATGCCA CTGTGTTTCA AGATGGTAGC AGCTTCCTCG AGCAGGAGGA ACGAAAAGCT 240
GTTTCTTTTC CACAGCCAGA TCTGCCTGAC AATCCCACAT ACTCAACAGA AGAAGAAAAA 300
CTGGCTTCAG ATGTTGGAGC AAATAAAAAA CAGGAAGGAC GTGTATTCGC AAACACTACT 360
TGGAGGGCCG GTACCTCCAA GGAAGTCTCC TTTGCAGTTG ATTTATGTGT ACTGTTCCCA 420
GAGCCAGCTC GTACCCATGA AGAGCAACAT AATTGTCCGG TCATAGGAGC AGGAAGTGTC 480
GACCTTGCGC CAGGATTTGG ACACCTCTGG AGCCAAACTG GATGTGGAAG CTCCAAAGGT 540
GCAGAAAAGG GGCCTCCAAA TGTGACTTT TACCTCTGTC CTGGAATCA CCTGACGCT 600
AGCTGTAGAG ATACTTACCA GTTTTCTGTC CCTGATTGGA CATGTGTAAC TTTAGCCACC 660
TACTCTGGGG GATCAACTAG ATCTTCAACT CTTTCCATAA GTCGTGTTC TCATCTTAAA 720
TTATGTACTA GAAAAAATTG TAATCCTCTT ACTATAACTG TCCATGACCC TAATGCAGCT 780
CAATGGTATT ATGGCATGTC ATGGGGATTA AGACTTTATA TCCCAGGATT TGATGTTGGG 840
ACTATGTTCA CCATCCAAA GAAATCTTT GTCTCATGGA GCTCCCCAA GCCAATCGGG 900
CCTTTAAGT ATCTAGGTGA CCCTATATTC CAGAAACACC CTGACAAAGT TGATTTAACT 960
GTTCTCTGCG CATCTCTAGT TCCTAGACCC CAGCTACAAC AACAACATCT TCAACCCAGC 1020
CTAATGTCTA TACTAGTTGG AGTACACCAT CTCCTTAACC TCACCCAGCC TAAACTAGCC 1080
CAAGATTGTT GGCATATGTT AAAAGCAAAA CCCCCTTATT ATGTAGGATT AGGAGTAGAA 1140
GCCACACTTA AACGTGGCCC TCTATCTTGT CATACACGAC CCCGTGCTCT CACAATAGGA 1200
GATGTGCTGT GAAATGCTTC CTGTCTGATT AGTACCGGGT ATAACCTATC TGCTTCTCCT 1260
TTTCAGGCTA CTTGTAATCA GTCCCTGCTT ACTTCCATAA GCACCTCAGT CTCTTACCAA 1320
GCACCCAACA ATACCTGGTT GGCCTGCACC TCAGGTCTCA CTGCTGTCAT TAATGGAAC 1380

```


5 GAACCAAGGAC CTCCTCTGTG CGTGTAGTTC CATGTACTTC CCCAGGTATA TGTGTACAGT 1440
 GGACCAGAAG GACGACAAC CATCGCTCCC CCTGAGTTAC ATCCCAGGTT GCACCAAGCT 1500
 GTCCCACATTC TGGTTCCCTT ATTGGCTGGT CTTAGCATAG CTGGATCAGC AGCCATTTGGT 1560
 ACGGCTGCCC TGGTTCAGG AGAACTGGA CTAATATCCC TGTCTCAACA GGTGGATGCT 1620
 GATTTTAGTA ACCTCCAGCT TGCCATAGAT ATACTACATT CCCAGGTAGA GTCTCTGGCT 1680
 GAAGTAGTTC TTCAAACTG CCGATGCTTA GATCTGCTAT TCCTCTCTCA AGGAGGTTTA 1740
 TGTGCGAGCTC TAGGAGAAAG TTGTGCTTC TATGCCAATC AATCTGGAGT CATAAAAGGT 1800
 ACAGTAAAAA AAGTTCGAGA AAATCTAGAT AGGCACCAAC AAGAACGAGA AAATAACATC 1860
 10 CCCTGGTATC AAAGCATGTT TAACTGGAAC CCATGGCTAA CTACTTTAAT CACTGGGTTA 1920
 GCTGGACCTC TCCTCATCCT ACTATTAAGT TTAATTTTGG GGCCTTGTAT ATTAAATTCG 1980
 TTTCTTAATT TTATAAAACA ACGCATAGCT TCTGTCAAAC TTACGTATCT TAAGACTCAA 2040
 TATGACACCC TTGTTAATAA CTGA

SEQ ID NO:58 PBJ7 Protein sequence

Protein Accession #: FGENSEH predicted

1 11 21 31 41 51
 | | | | |
 20 MDSCLOHMRD LLYLLQELRC LNPATLLPDP DSTTFVHDCQ DLLETTKTGQ PDLQDVPLEK 60
 ADATVFTDGS SFLQGERKA VSPQPDLPD NPYSTEEEEK LASDVGANKN QEGRVFANTT 120
 WRAGTSKEVS FAVDLCLVFP EPARTHEEQH NLPVIGAGSV DLAAFGHSG SQTGCGSSKG 180
 AEKGLQNVDF YLCPGNHEDA SCRDTYQFFC PDWTCVTLAT YSGGSTRSST LSI SRVPHPK 240
 LCTRKNCNPL TITVHDPNAA QWYVGMSWGL RLYIPGFDVG TMFTIQKIL VSWSSPKPIG 300
 25 PLTDLGDPFI QKHPDKVDLT VPLPFLVPRP QLQOHLQPS LMSILGGVHH LLNLTQPKLA 360
 QDCWLCLKAK PPIYVGLGVE ATLKRGLPSC HTRPRALTIG DVSGNASCLI STGYNLSASP 420
 FOATCNQSLT TSISTSVSYQ APNNTWLACT SGLTRCINGT EPGLLCVLV HVLPPQVYVYS 480
 GPEGRQLIAP PELHPRHLQA VPLLVLPLAG LSIAGSAAIG TAALVQGETG LISLSQQVDA 540
 DFNLSQSAID ILHSQVESLA EVVLQNCRL DLLFLSQGGL CAALGESCCF YANQSGVIKG 600
 30 TVKKVRENLD RHQQUERENNI PWYQSMFNWN PWTTLITLGL AGPLLLILLS LIFGPCILNS 660
 FLNFIKQRIA SVKLTLYLKTQ YDTLVNN

SEQ ID NO:59 PCQ1 DNA SEQUENCE

Nucleic Acid Accession #: NM_019005

Coding sequence: 182-1885 (underlined sequences correspond to start and stop codons)

35 1 11 21 31 41 51
 | | | | |
 40 TGATGGTGGGA AATTTCTTGA AACCGCTCTC GTAATTTGCC ACGTGTGTGT GCAAAATATTC 60
 TGGTGAATGA ACACAGAATC AGCATGGCTT TCCTTTGCTG AGAAATCACT GATGGGAAGT 120
 GAGACTTGTT AAACCTTGAAA GTGAATGGAC CTGAGTGGAC CCTTTGATCA CATCAGTAAA 180
 CATGAGCGGT ACCAAACCTG ATATTTTATG GGCACCACAC CATGTTGATA GATTTGTGTG 240
 GTGTGACTCA GAACTAAGTC TTTATCATGT GGAATCTACT GTGAATTCAG AACTCAAAGC 300
 TGGATCTTTA CGTTTATCTG AAGACTCTGC AGCTACATTA CTGTCAATAA ATTCTAGATAC 360
 45 ACCCTATATG AAATGTGTGG CCTGGTATCT TAATTATGAT CCTGAATGTC TGCTGGCAGT 420
 TGGACAAGCA AATGGTTCGAG TTGTACTTAC AAGCCTTGGT CAAGATCATA ACTCAAAGTT 480
 CAAAGATTTG ATAGGAAAAG AGTTTGTTC CAAACATGCA CGACATGTA ATACCTTTGC 540
 CTGGAATCCA CTGGATAGTA ACTGGCTAGC TGCTGGTTTA GATAAGCACA GAGCTGACTT 600
 TTCAGTGCTA ATATGGGATA TCTGCAGCAA ATATACCTCT GATATAGTTC CCATGGAAAA 660
 50 AGTGAAACTT TCAGCAGGTG AAACCTGAAAC AACATTATTA GTAACAAAAC CACTTTATGA 720
 GTTAGGACAG AATGATGCTT GTCTGTCTCT TTGTTGGCTT CCACGAGACC AGAAACTTCT 780
 CCTTGTGTGT ATGCATCGTA ACCTAGCTAT ATTGTATCTT CGGAATACAA GCCAAAAGAT 840
 GTTCGTAAT ACAAAAGCTG TTCAGGGTGT GACGGTAGAC CCATATTTTC ACGATCGTGT 900
 TGCTTCCTTC TATGAAGGTC AGGTTGCAAT ATGGGATCTT AGAAATTTG AGAAGCCAGT 960
 55 TTTGACATTG ACTGAGCAAC CAAAACCTT AACAAAAGTA GCATGGTGTG CCACTAGGAC 1020
 TGGTCTACTT GCCACTTTAA CAAGGGATAG TAATATTTAT AGATTGTATG ATATGCAGCA 1080
 TACACCCACT CCCATTGGGG ATGAACTGA ACCCACAATA ATTGAAGAA GTGTGCAACC 1140
 TTGTGACAA TACATTGCTT CCTTTGCGTG GCATCCAACA AGTCAAAATC GAATGATAGT 1200
 TGTAACFCCC AACCGAACAA TGTCAGACTT CACTGTTTTT GAAAGGATAT CTCTTGCTG 1260
 GAGCCCAATT ACATCTTTAA TGTGGGCTTG TGGTCGTCAT TTATATGAAT GTACGGAAGA 1320
 60 AGAAAATGAT AATCTCTTAG AAAAAGATAT AGCAACGAAG ATCGCTCTTC GGGCTTTATC 1380
 AAGGTATGGA CTTGATACAG AGCAGGTGTG GAGGAACCAC ATTTTAGCTG GAAATGAAGA 1440
 TCCACAGCTC AAGTCACTCT GGTATACTCT GCACCTTATG AAGCAATACA CAGAAGATAT 1500
 GGATCAGAAA TCTCCAGGCA ACAAGGATC ATTGGTTTAT GCAGGAATTA AATCAATTGT 1560
 65 AAAGTCATCG TTGGGAATGG TGGAAAGCAG CAGACATAAT TGGAGTGGGT TGGATAAGCA 1620
 AAGTGATATT CAAAACTTAA ATGAAGAGAG AATCTTAGCT TTACAGCTTT GTGGGTGGAT 1680
 AAAGAAAGGA ACGGATGTAG ACGTGGGGCC ATTTTGAAC TCCTTTGTAC AAGAAGGGGA 1740
 ATGGGAAAGA GCTGCTGCTG TGGCATTTGT CAACCTGGAT ATTGCGCCAG CAATCCAAAT 1800
 CCTGAATGAA GGGGCACTCT CTGAAAAGG CAGGAGATCT GAATCTCAAT GTGGTAGCAA 1860
 70 TGGCTTTATC GGGTTATACG GATGAGAAGA ACTCCCTTTG GAGAGAAATG TGTAGCACAC 1920
 TGCGATTACA GCTAAAATAAC CCGTATTGTG GTGTCATGTT TGCATTCTCT ACAAGTGAAA 1980
 CAGGATCTTA CGATGGAGTT TTGTATGAAA ACAAAAGTTC AGTACGTGAC AGAGTGGCAT 2040
 TTGCTGTGTA ATTCTTAGT GATACTCAGA TACATCGAAA AGTTGACCAA TGAATGAAA 2100
 GAGGCTGGAA ATTTGGAAGG AATTTTGCTT ACAGGCCCTA CTAAAGATGG AGTGGACTTA 2160
 ATGGAGAGTT ATGTTGATAG AACTGGAGAT GTTCAACACG CAAGTTACTG TATGTTACAG 2220
 75 GGTTCACCTT TAGATGTTCT TAAAGATGAA AGGTTTCACT ACTGGATTGA GAATTATAGA 2280
 AATTTATTAG ATGCTGGAG GTTTTGGCAT AAACGAGCTG AATTTGATAT TCACAGGAGT 2340
 AAGTTGGATC CCAAGTCCAA GCCTTTAGCA CAAGTTTGTG TGAGTTGCAA TTCTGTGGC 2400
 AAGTCAATCT CCTACAGCTG TTCAGCTGTG CCTCATCAGG GCAGAGGTTT TAGTCAGTAT 2460
 80 GGTGTGAGTG GCTCACCAC GAAATCTAAA GTCACAAAGT GTCCCTGGCTG TCGAAAACCA 2520
 CTTCTCGAT GTCCGCTTTG TCTCATTAAT ATGGGAACAC CAGTTTCTAG CTGTCCTGGA 2580

GGAACCAAAAT CAGATGAAAA AGTGGACTTG AGCAAGGACA AAAAAATTAGC CCAATTAAAC 2640
 AACTGGTTTAA CATGGTGTCAT TAATTGCAGG CACGGTGGAC ATGCTGGACA TATGCTTAGT 2700
 TGGTTCAGGG ACCATGCAGA GTGCCCTGTG TCTGCATGCA CGTGTAATG TATGCAGTTG 2760
 GATACAACGG GGAATCTGGT ACCTGCAGAG ACTGTCCAGC CATAAAATGT TACCACCTTA 2820
 AGAGAACCCT TCAAGTGTGG AGCTTCTAG TAGGTGTCTT TCATAGCTCA GAAACATACC 2880
 TCAGAACCAAG CCAATCATGA CTTACCTGTA ATGGGAAAAT AAATCATCTT ATCAGAAAAA 2940
 AAAAAAAAAA AAAAAAAAAA

SEQ ID NO:60 PCQ1 Protein sequence

Protein Accession #: NP_061878

1 11 21 31 41 51
 MSGTKPDILW APHVHDFRVV CDELSLYHV ESTVNSLKA GSLRLSEDSA ATLLSINSDT 60
 PYMKCVAWYL NYDPECLLAV GQANGRVVLT SLGQDHNKSF KDLIGKEFVP KHARQCNTLA 120
 WNPLDSNWL AGLDKHRADF SVLIWDICSK YTPDIVPMEK VKLSAGETET TLLVTKPLYE 180
 LGQNDACLSL CWLPRDQKLL LAGMHRNLAI FDLRNTSQKM FVNTKAVQGV TVDPYFHDRV 240
 ASFYEGQVAI WDLRKFEKPV LTLTEQPKPL TKVAWCPTRT GLLATLTRDS NIIRLYDMQH 300
 TPTPIGDETE PTIIRSVQVP CDNYIASFAW HPTSQNRNIV VTPNRMTSDF TVFERISLAW 360
 SPITSLMWAC GRHLYECTEE ENDNSLEKDI ATKMLRLALS RYGLDTEQVW RNHILAGNED 420
 POLKSLWYTL HFMKQYTEDM DQKSPGNKGS LUYAGIKSIV KSSLGMVESS RHNWSGLDKQ 480
 SDIQNLNEER ILALQLCGWI KKGTDVDVGP FLNSLVQEGE WERAAVALF NLDIRRAIQI 540
 LNEGASSEKG RRSESQCGSN GFILGLY

SEQ ID NO:61 PDG3 DNA SEQUENCE

Nucleic Acid Accession #: U42359

Coding sequence: 563-775 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 TTGTACATCT TAACAACCTT AAGCTGTACA AATAGANCAA TAATATCTAA ATGGTGTGAT 60
 GATCAGCCCA CAGTACACAT CATTGATGAG AATTTCACCTG GTCTCAACCT TTCTCATGCT 120
 GAGTCTCTGGC TTGTGTAATAT GACTTATAAA GGTCCAAGGA TTTAGAGATG ATTAAGAGAT 180
 AAGCTGGCAT TCTGTAAAGG CACCATCGTC TATCCCTGT CTTATCTAGA TAAAGAATGT 240
 AGTGCTAAAT CTGTGTAATA TATTGTACAA ATGGAATTC AATCTTAAGG ATTATTTTTT 300
 CCATATTGTT GTATTTCATG GTGGTGTATT GGAAAGTGAT CTGGACTTTG AGTGAGAAGA 360
 TGTGATTGTT ACCATGGCAC TTAATAACTC TATAACCTCA GGCAAGTCTT TTAATCTTCT 420
 CTGAGCTCA GTTTTCTCTCA TTTTCAAAAT ATAGAGAGTA TAACATTTAT CTCATAAGAC 480
 AAGTTGTAGT AAATTACTGT TTTACAAATG TAAGATAACT TTTAACTGTG AGATTCCATA 540
 TTCCAGTCTT ACATTATTAT GTTTATCTGC CACAGGGAGA AGTCTCTAGA TAAAAATGTC 600
 TACCAAAAGA CTGACACGTT GAGTTAATCA TTTGACAGAT GCAAATGCTT CCACCCCAAA 660
 CAAATATACT TTCTTTAACT TCTGTGTGGG TATCACTTAG GGAATAAAG GCAGGCAACA 720
 AAATATTTTT TAATTCATATC TTAGGAAAAA TTGTAGNCAA ATCTTTTTTT CCCATTAAACA 780
 AATAATGTAA CCTTAATAT TCAAGGGGTA ATAAAAATAC AAAGTCTTCC AAACAGGTAA 840
 CTACTTGAA AACTTT

SEQ ID NO:62 PDG3 Protein sequence

Protein Accession #: AAB18375

1 11 21 31 41 51
 MGARGAPSR RQAGRRRLRYL PTGSFPFLLL LLLLCIQLGG GQKKKENLLA EKVEQLMEWS 60
 SRRSIFRMNG DKFRFKIKAP PRNYSMIVMF TALQPQRQCS VCRQANEEOY ILANSWRYSS 120
 AFCNKLFFSM VDYDEGTDVF QQLNMNSAPT FXHXPPKGRP KRADTFDLQR IGFAEQBLAK 180
 WIADRTDVHI RVFRPPNYSG TIALALLVSL VGGLLYXRRN NLEFIYNTKG WAMVSLCIVF 240
 AMTSQGMWNH IRGPPYAHKN PHNGQVSYIH GSSQAQFVAE SHIILVLNAA ITMGMVLLNE 300
 AATSKGDVGK RRIICLVGLG LUVFFSFLL SIFRSKYHGY PYSDDLDFE

SEQ ID NO:63 PDG8 DNA SEQUENCE

Nucleic Acid Accession #: AL080235

Coding sequence: 245-453 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 GGTGCGCGCA CCGGCCCGCT CCGGCCCGCC GCCGCCCCCA GCGCCGCCGC CGCCACCGCC 60
 GGGGCGGCCA CCGCGCTGCC AGCCTACCCC GCGGCCGAGC CGCCCGGGCC GCTGTGGCTG 120
 CAGGGCGAGC CGCTGCATTG CTGCTGCCCTA GACTTCAGCC TGGAGGAGCT GCAGGGCGAG 180
 CCGGGCTGGC GGCTGAACCG TAAGCCCATTT GAGTCCACGC TGGTGGCCTG CTTCATGACC 240
 CTGGTCATCG TGGTGTGGAG CGTGGCCGCC CTCATCTGGC CGGTGCCCAT CATCGCCGCC 300
 TTCTGCCCCA ACGGCATGGA ACAGCGCCGG ACCACCGCCA GCACCAACGC AGCCACCCCC 360
 GCCCAGTGC CCGCAGGGAC CACCCAGCC GCCCGCGCCG CCGCCGCTGC CGCCCGCGCC 420
 GCGGCCGTCA CTTCGGGGGT GCGGACCAAG TGACCCGCTC CGCTCCTCCC TGTGTCCGTC 480
 CTGTGTCCGC GCGCGGGGT GCCTTTCCCG CCGGGGACTC GGCCGTTGTC CTTCGTGCTG 540
 TAGTTATCGT TAGTTCTCTT TCCCGAGATG GGGCCGCCGA GAGACCCAG CGCCTTTGAA 600
 AAGCAAGGTT TGTGCTGCGC TTCCAGTTCC GAAAAGCAGA TGTTTAAGCC CTGTGACTGA 660
 GGGTGGGATC GCAGCTCCGA AGACGGAGAG GAGGGAAATG GGGCCCTTTC CCCTCTATTG 720
 CATCCCCCTG CCGGACTCCT TCCCGCACC CACGTGCCCT AGATTTCATG CAGAAAATGA 780
 CCAAAATCCTG TGTATTGTTT TTATATATTT AATAACTGTT TTAATGAAA GTTTTAGTAA 840
 AAAAAATACA AAACAAAAG ATTAATATGC TATTGCTGTA GTAAGAGAA CTCTTTGTAT 900
 CTGAACATAG TTGTATTGTA AATTGTGTGT TTTTAAATTT ATTTAAAAAT GGGGGGAGGG 960

CATGGGAAGG ATTTAACACC GATATATTGT TACCGCTGAA AATGAACTTT ATGAACCTTT 1020
 TCCAAGTTGA TCTATCCAGT GACGTGGCCT GGTGGGCGTT TCTTCTTGTA CTTATGTGGT 1080
 TTTTGGCTT TTAATACAGA CATTTTCCTC CAAAAAAGG AAAAAAAGG

5

SEQ ID NO:64 PDG8 Protein sequence
 Protein Accession #: CAB45781

10 1 11 21 31 41 51
 | | | | |
 GRRTGRLRPA AAPSAAAATA GAPTALPAYP AAEPPGPLWL QGEPLHFCCL DFSLEELQGE 60
 FGWRLNRKPI ESTLVACFMT LVIVVWSVAA LIWVVPPIAG FLENGMEQRR TTASTTAATP 120
 AAVPAGTTAA AAAAAAAAAA AAVTSGVATK

SEQ ID NO:65 PDM1 DNA SEQUENCE

15

Nucleic Acid Accession #: NM_006765
 Coding sequence: 149-1195 (underlined sequences correspond to start and stop codons)

20 1 11 21 31 41 51
 | | | | |
 CGGCCGCGGC CGGGTCCCT CGCAAGCCG CTGCCATCCC GGAGGGCCCA GCCAGCGGGC 60
 TCCCGGAGGC TGGCCGGGCA GCGGTGGTGC GCGGTAGGAG CTGGGCGCGC ACGGCTACCG 120
 CGCGTGGAGG AGACACTGCC CTGCCCGGAT GGGGGCCCGG GCGGCTCCTT CACGCGGTAG 180
 GCAAGCGGGG CGGCGGCTGC GGTACCTGCC CACCGGGAGC TTTCCTTCC TTCTCCTGCT 240
 25 GCTGCTGCTC TGCATCCAGC TCGGGGAGG ACAGAAGAAA AAGGAGAATC TTTTAGCTGA 300
 AAAAGTAGAG CAGCTGATGG AATGGAGTTC CAGACGCTCA ATCTTCCGAA TGAATGGTGA 360
 TAAATTCGGA AAATTATATA AGGCACCACC TCGAAACTAT TCCATGATTTG TTATGTTTAC 420
 TGCTCTTCAG CCTCAGCGGC AGTGTCTGT GTGCAGGCAA GCTAATGAAG AATATCAAAT 480
 ACTGGCGAAC TCCTGGCGCT ATTCATCTGC TTTTGTAAAC AAGCTCTTCT TCAGTATGGT 540
 30 GGAATATGAT GAGGGGACAG ACGTTTTC CAAGCTCAAC ATGAACCTCTG CTCCTACATT 600
 CAYGCATTTW CTTCCAAAAG GCAGACCTAA GAGAGCTGAT ACTTTTGACC TCCAAAGAAT 660
 TGGATTTGCA GCTGAGCAAC TAGCAAAGTG GATTGCTGAC AGAACGGATG TTCATATTCG 720
 GGTTTTCAGA CCACCCAAC ACTCTGGTAC CATTTGCTTTG GCCCTGTAG TGTGCTTGT 780
 TGAGAGTTTG CTTATATNGA GAAGGAACAA CTGGAGTTTC ATCTATAACA AGACTGGTTG 840
 35 GGCATGGTG TCTCTGTGTA TAGTCTTTGC TATGACTTCT GGCCAGATGT GGAACCATAT 900
 CCGTGGACCT CCATATGCTC ATAAGAACCC ACACAATGGA CAAGTGAGCT ACATTCATGG 960
 GAGCAGCCAG GCTCAGTTTG TGGCAGAATC ACACATTAT CTGGTACTGA ATGCCGCTAT 1020
 CACCATGGGG ATGGTTCTTC TAAATGAAGC AGCAACTTCG AAAGGCGATG TTGGA AAAAG 1080
 40 ACGGATAATT TGCCTAGTGG GATTGGGCTT GGTGGTCTTC TTCTTCAGTT TTCTACTTTC 1140
 AATATTTTCGT TCCAAGTACC ACGGCTATCC TTATAGTGAT CTGGACTTTG AGTGAGAAGA 1200
 TGTGATTTGG ACCATGGCAC TTAAAACTC TATAACCTCA GCTTTTAAAT TAAATGAAGC 1260
 CAAGTGGGAT TTGCATAAAG TGAATGTTTA CCATGAAGAT AAACGTGTCC TGACTTTTATA 1320
 CTATTTTGAA TTCAATCAAT TCAATGTGAT CAGCTAGCTT ATTCTGTGTG ACTTTTPTTA 1380
 45 AACTGTGGGT TTTCTTAGTA AATTTAATTT ACAGAAATCA ATGGTAGCAT TTAGTAATCT 1440
 ACAAAGGAAA TATCAAAGTG TTTTCAAGC CTGTTATATY CAGTGTGTC CACAGGATG 1500
 CAATAATGA CAATGTAATT A

50

SEQ ID NO:66 PDM1 Protein sequence:
 Protein Accession #: NP_006756

55 1 11 21 31 41 51
 | | | | |
 MGARGAPSR RQAGRRRLYL PTGSFFPFL LLLLCIQLGG GQKKKENLLA EKVEQLMEWS 60
 SRRSIFRMNG DKFRKFIPAP PRNYSIMVMF TALQPORQCS VCRQANEVEQ ILANSWRVYS 120
 AFCNKLFFSM VDYDEGTDVF QQLNMNSAPT FXHXPCKGRP KRADTFDLQR IGFAAEQLAK 180
 WIADRTDVHI RVFRPPNYSG TIALALLVSL VGGLLYXRRN NLEFIYNKTG WAMVSLCIVF 240
 AMTSGQMWNH IRGPPYAHKN PHNGQVSYIH GSSQAQFVAE SHIILVLNAA ITMGVLLNE 300
 AATSKGDVKG RRIICLVGLG LVVFFFSFLL SIFRSKYHGY PYSDLDFE

60

SEQ ID NO:67 PDM2 DNA SEQUENCE

65

Nucleic Acid Accession #: NM_000947
 Coding sequence: 88-1617 (underlined sequences correspond to start and stop codons)

70 1 11 21 31 41 51
 | | | | |
 GGTTCATAT GAACCTCTCC GCCACCGGG AACAGCTGGC TGCCACCGTT TGTGTTTCC 60
 GAGTTGTAT TCTTGCAAGT GACCAAGATG GAGTTTCTG GAAGAAAGCG GAGGAAGCTG 120
 AGGTGGCAG GTGACCAGAG GAATGCTTCC TACCCCTCAT GCCTTCAGTT TACTTGCAG 180
 CCACCTCTG AAAACATATC TTTAACAGAA TTTGAAAAC TGGCTATTGA TAGAGTTAAA 240
 TTGTTAAAT CAGTTGAAAA TCTTGAGAGT AGCTATGTGA AAGGAACTGA ACAATACCAG 300
 75 AGTAAGTTGG AGAGTGAGCT TCGGAAGCTC AAGTTTCTCT ACAGAGAGAA GCTAGAAGAT 360
 GAATATGAAC CACGAAGAAG AGATCATATT TCTCATTTTA TTTTGGCGCT TGCTTATTGC 420
 CAGTCTGAAG AACTTAGACG CTGGTTCATT CAACAAGAAA TGGATCTCCT TCGATTTAGA 480
 TTTAGTATTT TACCCAAAGG TAAATTCAG GATTTCTTAA AGGATAGCCA ATTGCAGTTT 540
 GAGGCTATAA GTGATGAAGA GAAGACTCTT CGAGAACAGG AGATTGTTGC CTCATCACCA 600
 80 AGTTTAAAGT GACTTAAAGT GGGGTCGAG TCATTATATA AGATCCCTTT TGCTGATGCT 660
 CTGGATTGTG TTCGAGGAAG GAAAGTCTAT TTGGAAGATG GCTTTGCTTA CGTACCACCT 720

	AAGGACATTG	TGGCAATCAT	CCTGAATGAA	TTTAGAGCCA	AACTGTCCAA	GGCTTTGGCA	780
	TTAACAGCCA	GGTCCTTGCC	TGCTGTGCAG	TCTGATGAAA	GACTTCAGCC	TCTGCTCAAT	840
	CACCTCAGTC	ATTCCCTACAC	TGGCCAAGAT	TACAGTACCC	AGGGAAATGT	TGGGAAGATT	900
	TCCTTAGATC	AGATTGATTT	GCTTCTTACC	AAATCCTTCC	CACCTTGCCAT	GCGTCAGTTA	960
5	CATAAAGCCT	TGCGGGAAAA	TCACCATCTT	CGTCATGGAG	GCCGAATGCA	GTATGGCCTA	1020
	TTTCTGAAGG	GCAATGGTTT	AACTTTGGAA	CAGGCATGTC	AGTCTTGGAA	GCAAGAATTT	1080
	ATCAAAAGAA	AGATGGATCC	AGACAAGTTT	GATAAAGGTT	ACTCTTACAA	CATCCGCTAC	1140
	AGCTTTGGAA	AGGAAGGCAA	GAGGACAGAC	TATACACCTT	TCAGTTGCCT	GAAGATTATT	1200
10	CTGTCCAATC	CACCAAGCCA	AGGGGATTAT	CATGGGTGCC	CATTCGGTCA	CAGTGATCCA	1260
	GAGCTGCTGA	AGCAAAAGTT	GCAGTCATAC	AAGATCTCTC	CTGGAGGGAT	AAGCCAGATT	1320
	TTGGATTAG	TAAAGGGGAC	ACATTACCAG	GTAGCCTGTC	AAAAATACTT	TGAGATGATA	1380
	CACAATGTGG	ATGATTTGGG	CTTTCTTTTG	AATCATCCTA	ATCAGTTCTT	TTGTGAGAGC	1440
	CAACGTATTCT	TAAATGGTGG	TAAAGACATA	AAGAAGGAAC	CTATCCAACC	AGAAACTCCT	1500
15	CAACCCAAAC	CAAGTGTCCT	GAAAACCAAG	GATGCATCAT	CTGCTCTGGC	CTCTTTAAAT	1560
	TCCTCTCTGG	AAATGGATAT	GGAAGGACTA	GAAGATTACT	TTAGTGGAAG	TTCTTAGGCA	1620
	GTTTATAAAC	CCCTTTTCCCT	CAATAGCCTG	TTTCCCTGTTT	TTAAGATTTT	GCCTTTGTTG	1680
	TTGAAAAGG	GTTCCTACTG	CACCAAGGCT	TAGTGCACTG	ACACAATTAC	AGCTGATTGC	1740
	AGCCTTGACC	TTCCAGCTC	AAGTGATCCT	CCTACCTCAG	CCTCCCAAGT	AGTTAGGACA	1800
20	CACAGGTGTG	CACCTCATAT	CCAGATAAAT	TTTTTCAATT	TTTTTTTGTA	GAGGTGGGGG	1860
	GTCTCCCTAT	GTTCGCCAGG	CAGATCTCAG	ACTCCTGGGC	TCAGGCGATC	CTCACACCTC	1920
	AGCGTCCGAC	AGTGCTGGGA	TTACAGTTGT	GAGCCACTGT	GCCTGGCCTT	TTTTTTTTTT	1980
	TAACTTTTTC	GTTTAACTTC	TCTCTTCACT	GCATCCCAAT	CCATCTACAG	GCATGCACAC	2040
	TTATTAGGAA	AGGAGGTTTG	AGGTAACAAC	AGAGACTTTC	ACTATATTTT	GCTTTGACAG	2100
25	AAGGAAAGAG	GAGGAGTTTC	TATTAATAATC	TGTCACCTGA	GTGATGTCAT	TTAAGTCCTA	2160
	TTTTAGGAGA	TAAAAACAGC	TTTGGGGACT	GGTTAAAGTC	CCCCAGAAAC	TACAATAAAG	2220
	AACAACTTTT	GTTTTAACTC	TTAATCACTT	TGTAATTTTG	ACTCAATCCT	TTTCTGGACC	2280
	ATTTTGTGTA	ATAAATATCA	AAGTGT				

30 SEQ ID NO:68 PDM2 Protein sequence:

Protein Accession #: NP_000938

	1	11	21	31	41	51	
35	MEFSGRRKRR	LRLAGDQRNA	SYPHCLQFYL	QPPSENISLT	EFENLAIDRV	KLLKSVENLG	60
	VSVVKGTEQY	QSKLESBLRK	LKFSYREKLE	DEYEPRRRDH	ISHFILRLAY	CQSEELRRWF	120
	IQQEMDLLRF	RFSLLPKDKI	QDLKDSQLQ	FEAISDEEKT	LREQEIVASS	PSLSGLKLGF	180
	ESIYKIPFAD	ALDLFRGRKV	YLEDGFAYVP	LKDIVAIILN	EFRAKLSKAL	ALTARSLPAV	240
40	QSDERLQPLL	NHLSHSYTQG	DYSTQGNVKG	ISLDQIDLLS	TKSFPPCMRQ	LHKALRENHH	300
	LRHGGRMQYG	LFLKIGLTL	EQALQFWKQE	FIKGMDDPK	FDKGYSYNIR	HSFGKEGKRT	360
	DYTFFSCLKI	ILSNPPSQGD	YHGCPFRHSD	PELLKQKLQS	YKISPGGISQ	ILDVLKGYTHY	420
	QVACQKYFEM	IHNVDGCGFS	LNHPNQFFCE	SQRILNGGKD	IKKEPIQPET	PQPKPSVQKT	480
	KDASSALASL	NSSLEMDMEG	LEDYFSEDS				

45 SEQ ID NO:69 PDM3 DNA SEQUENCE

Nucleic Acid Accession #: NM_024840

Coding sequence: 108-491 (underlined sequences correspond to start and stop codons)

	1	11	21	31	41	51	
50	AATTCATACA	GGAGAGAAGT	CATATATATG	CAGTGATTGT	GGAAAAGGCT	TCATCAAGAA	60
	GTCTCGGCTC	ATTAATCATC	AGAGAGTTCA	TACAGGAGAG	AAACCACATG	GATGCAGCCT	120
	GTGTGGGAAG	GCCCTCTCCA	AAAGGTCAG	GCTCACTGAA	CACCAGAGAA	CTCATACAGG	180
	AGAGAAAGCC	TATGAATGCA	CTGAATGTGA	CAAAGCATTC	CGCTGGAAAT	CACAGCTCAA	240
55	TGCACATCAG	AAAGCTCACA	CAGGAGAGAA	GTCAATATAT	TGCCGTGATT	GTGGAAAAGG	300
	CTTCATTCAG	AAGGGAAATC	TCATTGTACA	TCAGCGAATT	CATACTGGAG	AAAAACCCCTA	360
	TATATGCAAT	GAATGTGGAA	AAGGCTTCAT	CCAAAAGGGC	AACCTCCTTA	TTTATCGACG	420
	TACTCACACT	GGAGAGAAAC	CCTATGAATG	CAATGAATGT	GGGAAAGGCT	TCAGCCAGAA	480
60	GACATGTTTA	ATATCCCATC	AGAGATTTCA	CACAGGAAAG	ACACCCTTTG	TATGTACTGA	540
	GTGTGGAAAA	TCCTGCTCAC	ACAAGTCAGG	TCTCATTAAC	CACCAGAGAA	TTACACACAGG	600
	AGAGAAACCC	TATACATGCA	GTGACTGTGG	GAAAGCTTTC	AGAGATAAAT	CATGTCTCAA	660
	CAGACATCGG	AGAACTCATA	CAGGGGAGAG	ACCGTATGGA	TGCTCTGATT	GTGGGAAAGC	720
	TTTCTCCCAC	TTGTCTATGCC	TTGTTTATCA	TAAGGGAATG	CTGCATGCAA	GAGAGAAATG	780
65	TGTAGGTTCA	GTCAAATTGG	AAAATCCTTG	CTCAGAGAGT	CATAGCTTAT	CACATACACG	840
	TGATCTCATC	CAGGATAAAG	ACTCTGTTAA	CATGGTGACT	CTGCAGATGC	CTTCTGTGGC	900
	AGCTCAGACC	CATTAACTA	ACAGTGCGTT	CCAAGCAGAG	AGCAAAGTAG	CCATTGTGAG	960
	CCAGCCTGTT	GCCAGAAGTT	CAGTCTCAGC	AGATAGTAGA	ATTTGCACAG	AATAAAAAACC	1020
	ATATGAATGC	AGTGAATGTG	GTAGTGCTTT	CAGTGATCAA	TTACATCATA	TGTCACAAAA	1080
70	AACACAGAGG	AACAACTGA	TATATTCAG	GTGGAAAGCC	CTTGAATAAA	ACCTTATGGC	1140
	TAATAAGCAT	ATACTCAGAG	AAAAATAGTA	TGAAGTGGAG	ACTGGGAAAT	TCTTTTATGG	1200
	GAAGATAGAT	CTTCTCATCA	GTGACCATAG	ATCACATCTT	CAGTGAGCTT	ATAGTTGGTA	1260
	GAAATATAAT	GATCATGGAA	AAGTCCTTGT	TCAGAAACAG	TACGCCAGTA	GGTATCAGGG	1320
	GGTTTACACA	GGAGAGAAAC	TTTTGGAAGA	CCTTTGAAGG	CTATGAATGT	GGCAGGGTTG	1380
75	CTAGTGGTAC	ATTCCTCCTT	ATCCTCAGAG	GGAATCATAT	AGAAATAAAA	CTATGAAAT	1440
	GTAACTAGAA	CATCTTTCAT	AAAATATGAA	AGAACACACG	AAGCAAATAA	GCCCTGTGAA	1500
	AAGGAGTATT	TTAGAGATTT	CGATCAGAAA	TCTAACATCA	TTATATGGCA	GATAATATAC	1560
	AGGATGTGTA	TTTTAGGACA	ATATACCTTG	AATCACTAGT	TGATATGTCA	ATGACTAATT	1620
	AAAAGGGGTT	GTACAGTTTA	CACATCATTG	GTAAATTTTA	TAGCACAATG	TACCTCTTCC	1680
80	CCCTTTTGTG	ATAAGAGTCT	TCTATTCCCA	ACCAAGATCA	TTATATGATT	AGCTCTTGTT	1740
	TTTCTTTGAT	TCCAAATTTT	TTCACTTGTT	ATTTCAAGCT	ACTGAAGCTC	TTCAAAGGA	1800

AAAATGTATT TAATTTAATA ATGTAACACA ACAAGTTTGG ATGTGTTTAA CTTTATAAAT 1860
AATCACCCCA GAGGAATGAA GTTCAAACCT TGTGAATAAC C

5

SEQ ID NO:70 PDM3 Protein sequence:

Protein Accession #: NP_079116

1	11	21	31	41	51	
MDAACVGRPS	PKGPGSLNTR	ELIQERSPMN	ALNVTKHSAG	NHSSMHIRKL	TQERSHIYAV	60
IVEKASFRRE	ISLYISEFIL	EKNPIYAMNV	EKASSKRATS	LFIDVLTLEL	NPMNAMNVGK	120
ASARRHV						

15

Nucleic Acid Accession #: NM_018455

Coding sequence: 341-955 (underlined sequences correspond to start and stop codons)

SEQ ID NO:71 PDM8 DNA SEQUENCE

1	11	21	31	41	51	
AATTTCGGCA	CGGGGGGGGAG	GCACAGTGAG	TCCACTGGGG	CACGGCAGCG	TCTAAGCCAC	60
AAGCCGACTG	ACATAAGCCA	GGTCCTAACG	GAGCCTATGT	GTAAGTCCAC	TACTGGTGCA	120
AGGTTGCACA	CTTCTAAGAA	GAGCGGCGTG	GGGGGCTCGG	CGACCTTCGC	TTCAGTCGCT	180
CCCCCGTGCA	GTCCCTCTGT	CCCAAGACAC	AGCCTGATGC	TTGTGCTCCG	GTGGGCGGAC	240
TTGGAGGCGG	CGGGAACATG	AATTGGTGGC	TTTGAAGGGC	GGCGAGCGGG	AACAGCTCTT	300
GAGGAGTGAG	ACTGCAGGAG	ATGTGGGCGG	TGCCAAAGAG	ATGGATGAGA	CTGTTGCTGA	360
GTTTCATCAAG	AGGACCACTT	TGAAAATCCC	CATGAATGAA	CTGACACAA	TCCTGAAGGC	420
CTGGGATTTT	TTGTCTGAAA	ATCAACTGCA	GACTGTAAAT	TTCCGACAGA	GAAAGGAATC	480
TGTAGTTTAC	CACTTGATCC	ATCTGTGTGA	GGAAAAGCGT	GCAAGTATCA	GTGATGCTGC	540
CCTGTTAGAC	ATCATTTTATA	TGCAATTTC	TCAGCACCAG	AAAGTTTGGG	ATGTTTTC	600
GATGAGTAAA	GGACCAGGTG	AAGATGTTGA	CCTTTTGTAT	ATGAAACAAT	TTAAAAATTC	660
GTTCAAGAAA	ATTCTTCAGA	GAGCATTA	AAATGTGACA	GTGAGCTTCA	GAGAAACTGA	720
GGAGAATGCA	GTCTGGATTC	GAATTGCCGT	GGGAACACAG	TACACAAAGC	CAAACCAAGT	780
CAAACCTACC	TACGTGGTGT	ACTACTCCCA	GACTCCGTAC	GCCTTCACGT	CCTCCTCCAT	840
GCTGAGGCGC	AATACACCGC	TTCTGGGTCA	GGAGTTAGAA	GCTACTGGGA	AAATCTACCT	900
CCGACAAGAG	GAGATCAATT	TAGATATTAC	CGAAATGAAG	AAAGCTTGCA	ATTAGTGAAC	960
ATGAAAGGAA	AATAAAAAAT	CCTCACAGTC	AAAAAAAAAA	AAAAA		

SEQ ID NO:72 PDM8 Protein sequence:

Protein Accession #: NP_060925

1	11	21	31	41	51	
MDETVAEFIK	RTILKIPMNE	LTILKAWDF	LSENQLQTVN	FRQRKESVVQ	HLIHLCEEKR	60
ASISDAALLD	IIYMQFHQHQ	KVWDVFQMSK	GPGEVDLFD	MKQFKNSFKK	ILQRALKNVT	120
VSFRETEENA	VWIRIAWGTQ	YTKPNQYKPT	YVVYYSQTPY	AFTSSSMLRR	NTPLLGQELE	180
ATGKIYLRQE	BIILDITEMK	KACN				

SEQ ID NO:73 PDM9 DNA SEQUENCE

Nucleic Acid Accession #: NM_016192

Coding sequence: 1-1125 (underlined sequences correspond to start and stop codons)

1	11	21	31	41	51	
ATGGTGCTGT	GGGAGTCCCC	GCGGCAGTGC	AGCAGCTGGA	CACTTTGCGA	GGGCTTTTGC	60
TGGCTGCTGC	TGCTGCCCGT	CATGCTACTC	ATCGTAGCCC	GCCCGGTGAA	GCTCGCTGCT	120
TTCCCTACCT	CCTTAAGTGA	CTGCCAAACG	CCCACCGGCT	GGAATGTGCT	TGTTATGAT	180
GACAGAGAAA	ATGATCTCTT	CCTCTGTGAC	ACCAACACCT	GTAATTTTGA	TGGGGAATGT	240
TTAAGAATTG	GAGACACTGT	GACTTGCGTC	TGTCAGTTCA	AGTGCAACAA	TGACTATGTG	300
CCTGTGTGTG	GCTCCAATGG	GGAGAGCTAC	CAGAATGAGT	GTTACCTGCG	ACAGGCTGCA	360
TGCAACACGC	AGAGTGAGAT	ACTTGTGGTG	TCAGAAGGAT	CATGTGCCAC	AGATGCAGGA	420
TCAGGATCTG	GAGATGGAGT	CCATGAAGGC	TCTGGAGAAA	CTAGTCAAAA	GGAGACATCC	480
ACCTGTGATA	TTTGCCAGTT	TGGTGACGAA	TGTGACGAAG	ATGCCGAGGA	TGCTGTGGT	540
GTGTGTAATA	TTGACTGTTC	TCAAACCAAC	TTCAATCCCC	TCTGCCTTTC	TGATGGGAAA	600
TCTTATGATA	ATGCATGCCA	AATCAAGAAA	GCATCGTGTC	AGAAACAGGA	GAAAATTGAA	660
GTCATGTCTT	TGGGTCGATG	TCAAGATAAC	ACAACTACAA	CTACTAAGTC	TGAAGATGGG	720
CATTATGCAA	GAACAGATTA	TGCAGAGAAT	GCTAACAAAT	TAGAAGAAAG	TGCCAGAGAA	780
CACCACATAC	CTTGTCCGGA	ACATTACAAT	GGCTTCTGCA	TGCATGGGAA	GTGTGAGCAT	840
TCTATCAATA	TGCAGGAGCC	ATCTTGCAGG	TGTGATGCTG	GTTTACTTGG	ACAACACTGT	900
GAAAAAAGG	ACTACAGTGT	TCTATACGTT	GTTCCCGGTC	CTGTACGATT	TCAGTATGTC	960
TTAATCGCAG	CTGTGATTGG	AACAATTTCAG	ATTGCTGTCA	TCTGTGTGGT	GGTCCCTGTC	1020
ATCAACAAGG	AATGCCCCAG	AAGCAACAGA	ATTACACAGC	AGAAGCAAAA	TACAGGGCAC	1080
TACAGTTTCA	ACAATACAAC	AAGAGCGTCC	ACGAGGTTAA	TCTGA		

SEQ ID NO:74 PDM9 Protein sequence:
Protein Accession #: NP_057276

```

5      1      11      21      31      41      51
      1 MVLWESPRQC SSWTLCEGFC WLLLLPVMLL IVARPVKLAA FPTSLSDCQT PTGWNC SGYD 60
      61 DRENDLFLCD TNTCKFDGEC LRIGDTVTCV CQFKCNNDYV FVCGSNGESY QNECYLRQAA 120
      121 CKQQSEILVV SEGSCATDAG SSGSDGVHEG SGETSQKETS TCDICQFGAE CDEDAEDVWC 180
      181 VCNIDCSQTN FNPLCASDGI SYDNACQIKE ASCQKQEKIE VMSLGRQDN TTTTTSKSDG 240
10     241 HYARTDYAEN ANKLEESARE HHIPCPHYN GFCMHGKCEH SINMQPESCR CDAGYTGQHC 300
      301 EKKDYSVLVY VPGFVRFPYV LIAAVIGTIQ IAVICVVVLC ITRKCPRSNR IHRQKQNTGH 360
      361 YSSDNTTRAS TRLI

```

SEQ ID NO:75 PDO1 DNA SEQUENCE

Nucleic Acid Accession #: NM_014324
Coding sequence: 89-1237 (underlined sequences correspond to start and stop codons)

```

20     1      11      21      31      41      51
      GGCGCCGGGA TTGGGAGGGC TTCTTGCAAG CTGCTGGGCT GGGGCTAAGG GCTGCTCAGT 60
      TTCTTTCAGC GGGGCACTGG GAAGCGCCAT GGCAGTGCAG GGCATCTCGG TCGTGGAGCT 120
      GTCCGGCCTG GCCCGGGGCC GTTCTGTGTC TATGGTCCGT GCTGACTTCG GGGCGCGTGT 180
      GGTACGCGTG GACCCGGCCG GCTCCCGCTA CGACGTGAGC CGCTTGGGCC GGGGCAAGCG 240
      CTCGCTAGTG CTGGACCTGA AGCAGCCGCG GGAGCCGCGT GCTGCGGCGT CTGTGCAAGC 300
25     GGTCGGATGT GGTGCTGGAG CCTTCCGCC CGGGTGTCTG GGAGAAACTC CAGCTGGGCC 360
      CAGAGATTCT GCAGCGGAAA AATCCAAGGC TTATTATATG CAGGCTGAGT GGATTGGGCC 420
      AGTTCAGGAA AGCTTCTGCC GGTAGCTGG CCACGATATC AACTATTTGG CTTTGTCAGG 480
      TGTCTCTCA AAAATTGGCA GAAGTGGTGA GAATCCGTAT GCCCGCTGA ATCTCGTGGC 540
      TGACTTGTCT GGTGGTGGCC TTATGTGTGC ACTGGGCATT ATAATGGCTC TTTTGCACCG 600
30     CACACGCACT GACAAGGGTC AGGTCATTGA TGCAAATATG GTGGAAGGAA CAGCATATTT 660
      AAGTCTCTTT CTGTGGAAAA CTCAGAAATC GAGTCTGTGG GAAGCACCTC GAGGACAGAA 720
      CATGTTGGAT GGTGGAGCAC CTTCCTATAC GACTTACAGG ACAGCAGATG GGGAAATCAT 780
      GGCTGTTGGA GCAATAGAAC CCCAGTCTTA CGAGCTGCTG ATCAAAGGAC TTGGACTAAA 840
      GTCTGATGAA CTTCCTCAATC AGATGAGCAC GGATGATTGG CCAGAAATGA AGAAGAAGTT 900
35     TGCAGATGTA TTGCAAAAGA AGACGAAGGC AGAGTGGTGT CAAATCTTTG ACGGCACAGA 960
      TGCTCTGTGT ATCTCCGTTG TGACTTTTGA GGAGGTTGTT CATCATGATC ACAACAAGGA 1020
      ACGGGGCTCG TTTATCACC A GTGAGGAGCA GGACGTGAGC CCCCCTCTTG CACCTCTGCT 1080
      GTTAAACACC CCAGCCATCC CTTCCTCCAA AGGGGATCCT TTCTATAGGAG AACACACTGA 1140
      GGAGATACTT GAAGAATTTG GATTACAGCC AGAAGAGATT TATCAGCTTA ACTCAGATAA 1200
40     AATCATTGAA AGTAATAAGG TAAAAGCTAG TCTCTAACTT CCAGGCCCAC GGCTCAAGTG 1260
      AATTGGAATA CTGCATTATC AGTGTAGAGT AACACATAAC ATTGTATGCA TGGAAACATG 1320
      GAGGAACAGT ATTACAGTGT CCTACCACCT TAATCAAGAA AAGAATTACA GACTCTGATT 1380
      CTACAGTGAT GATTGAATTC TAAAATGGT TATCATTAGG GCTTTTGATT TATAAACTT 1440
      TGGGTACTTA TACTAAATTA TGGTAGTTAT TCTGCCTTCC AGTTTGTCTG ATATATTTGT 1500
45     TGATATTAAG ATTCTTGACT TATATTTTGA ATGGGTTCTA GTGAAAAGG AATGATATAT 1560
      TCTTGAAGAC ATCGATATAC ATTATTTTAC ACTCTTGATT CTACAATGTA GAAAATGAGG 1620
      AAATGCCACA AATTGTATGG TGATAAAAGT CACGTGAAAC AGAGTGATTG GTTGCATCCA 1680
      GGCTTTTGT CTGGGTGTTT ATGATCTCCC TCTAAGCACA TTCCAAACTT TAGCAACAGT 1740
      TATCACACTT TGTAAATTTG AAAGAAAAGT TTCACCTGTA TTGAATCAGA ATGCCTTCAA 1800
50     CTGAAAAAAA CATATCCAAA ATAATGAGGA AATGTGTTGG CTCCTACAGT AGAGTCCAGA 1860
      GGGACAGTCA GTTTTAGGGT TGCTGTATC CAGTAAGTGG GGGCCTGTTT CCCCCTGGGT 1920
      CTCTGGGCTG TCAGCTTTCC TTTCTCCATG TGTGTTGATT CTCTCAGGC TGGTAGCAAG 1980
      TTCTGGATCT TATACCAAC ACACAGCAAC ATCCAGAAAT AAAGATCTCA GGACCCCCA 2040
      AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA

```

SEQ ID NO:76 PDO1 Protein sequence:
Protein Accession #: NP_055139

```

60     1      11      21      31      41      51
      1 MALQGISVVE LSGLAPGRXC AMVLADFGAR VVRVDRPGSR YDVSRLGRGK RSLVLDLKQP 60
      61 REPRAAASVQ AVGCAAGALP PRCHGETPAG PRDSAAGKSK AYLCQAEWIW PVQESFCRLA 120
      121 GHDINYLALS GVLKIGRSG ENPYAPLNLV ADFAGGGLMC ALGIIMALFD RTRTDKQOVI 180
      181 DANMVEGTAY LSSFLWKTKQ SSLWEAPRQO NMLDGGAPFY TTVRTADGEF MAVGAIEPQF 240
65     241 YELLIKGLGL KSDLPNQMS TDDWPEMKKK FADVFAKKTK AEWCQIFDGT DACVTPVLTF 300
      301 EEVVHHDHKN ERGSFITSEE QDVSPRLAPL LLNTPAIPSS KGDPFFIGEHT EEILEEFGFS 360
      361 REEIIQLNSD KIIESNKVKA SL

```

SEQ ID NO:77 PDO3 DNA SEQUENCE

Nucleic Acid Accession #: AB028951
Coding sequence: 97-1128 (underlined sequences correspond to start and stop codons)

```

75     1      11      21      31      41      51
      GTTAAATCCT TACTTTACCA GATTCTTGAT GGTATCCATT ACCTCCATGC AAATTGGGTG 60
      CTTCACAGAG ACTTGAAACC AGCAAAATATC CTAGTAATGG GAGAAGGTCC TGAGAGGGGG 120
      AGAGTCAAAA TAGCTGACAT GGGTTTGTCC AGATTATTCA ATTCTCCTCT AAAGCCACTA 180
      GCAGATTGGG ATCCAGTAGT TGTGACATTT TGGTATCGGG CTCAGAAACT TTTGCTTGGT 240
80     GCAAGGCATT ATACAAAGGC CATTGATATA TGGGCAATAG GTTGTATATT TGCTGAATTG 300
      TTGACTTCGG AACCTATTTT TCACTGTCGT CAGGAAGATA TAAAAACAAG CAATCCCTTT 360

```

	CATCATGATC	AACTGGATCG	GATATTTAGT	GTCATGGGGT	TTCCTGCAGA	TAAAGACTGG	420
	GAAGATATTA	GAAAGATGCC	AGAATATCCC	ACACTTCAAA	AAGACTTTAG	AAGAACAACG	480
	TATGCCAACA	GTAGCCTCAT	AAAGTACATG	GAGAAACACA	AGGTCAAGCC	TGACAGCAAA	540
5	GTGTTCCTCT	TGCTTCAGAA	ACTCCTGACC	ATGGATCCAA	CCAAGAGAAT	TACCTCGGAG	600
	CAAGCTCTGC	AGGATCCCTA	TTTTTCAGGAG	GACCCTTTGC	CAACATTAGA	TGTATTTGCC	660
	GGCTGCCAGA	TTCCATACCC	CAAACGAGAA	TTCTTTAATG	AAGATGATCC	TGAAAGAAAA	720
	GGTGACAAGA	ATCAGCAACA	GCAGCAGAAC	CAGCATCAGC	AGCCACACAG	CCCTCCACAG	780
	CAGGCAGCAG	CCCTCCACAA	GGCGCCCCCA	CCACAGCAGA	ACAGCACCCA	GACCAACGGG	840
10	ACCGCAGGTG	GGGCTGGGGC	CGGGGTCGGG	GGCACCAGGAG	CAGGGTTGCA	GCACAGCCAG	900
	GACTCCAGCC	TGAACCAAGT	GCCCTCCAAAC	AAGAAGCCAC	GGCTAGGGCC	TTCAGGCGCA	960
	AACTCAGGTG	GACCTGTGAT	GCCCTCGGAT	TATCAGCAGT	CCAGTTCTCG	CCTGAATTAC	1020
	CAAGCAGCGG	TTCAGGGATC	CTCTCAGTCC	CAGAGCACAC	TTGGCTACTC	TTCCCTCGTCT	1080
	CAGCAGAGCT	CACAGTACCA	CCCATCTCAC	CAGGCCACAC	GGTACTGACC	AGCTCCCGTT	1140
	GGGCCAGGCC	AGCCCAGCCC	AGAGCACAGG	CTCCAGCAAT	ATGTCTGCAT	TGAAAGAAC	1200
15	CAAAAAATG	CAAACTATGA	TGCCATTAA	AACTCATACA	CATGGGAGGA	AAACCTTATA	1260
	TACTGAGCAT	TGTGCAAGGAC	TGATAGCTCT	TCCTTATGTA	CTTAAAGAAG	ATCTTGTGTA	1320
	AGTTTCCCCA	GCACCCCTTC	CTTGCATGTG	TTCCATTGTG	ACTTCTCTGA	TAAAGCGTCT	1380
	GATCTAATCC	CAGCAGCTCT	GTAACCTTCA	GCATTTCTTT	GAAGGATTTT	CTGGTGCACC	1440
20	TTTCTCATGC	TGTAGCAATC	ACTATGGTTT	ATCTTTTCAA	AGCTCTTTTA	ATAGGATTTT	1500
	AATGTTTTAG	AAACAGGATT	CCAGTGGTGT	ATAGTTTTAT	ACTTCATGAA	CTGATTTAGC	1560
	AACACAGGTA	AAAATGCACC	TTTTAAAGCA	CTACGTTTTT	ACAGACAATA	ACTGTCTGTC	1620
	TCATGGAAGT	CTTAAACAGA	AACTGTTACT	GTCCCAAAGT	ACTTTACTAT	TACGTTCTGA	1680
	TTTATCTAGT	TTCAGGGGAG	GTCTAATAAA	AAGACAAGCG	GTGGGACAGA	GGGAACCTAC	1740
25	AACCAAAAA	TGCCTAGATC	TTTGCAGTTA	TGTGCTTTAT	GCCACGAAGA	ACTGAAGTAT	1800
	GTGGTAATTT	TTATAGAATC	ATTCAATATG	AACTGAGTTC	CCAGCATCAT	CTTATCTGTA	1860
	ATAGCATTTA	GTAATTAAGA	ATTACAAATTT	TAACCTTCAT	GTAGCTAAGT	CTACCTTAAA	1920
	AAGGGTTTCA	AGAGCTTTGT	ACAGTCTCGA	TGGCCACAC	CAAAACGCTG	AAGAGAGTAA	1980
	CAACTGCAGT	AGGATTTCTG	TAAAGAGTAA	TTTTGATCAA	AAGACGTGTT	ACTTCCCTTT	2040
30	GAAGGAAAA	TTTTTAGTGT	GATATTGTACA	TAAAGTCGGC	TTCTCTAAAG	AACCAATGGT	2100
	TTCTTCACAT	CTGGGCTGTC	GTGAGTAACT	TTCTTGCATA	ATCAAGGTTA	CTCAAGTAGA	2160
	AGCCTGAAAA	TTAATCTGCT	TTTAAATAAA	AGAGCAGTGT	TCTCCATTCT	TATTTGTATT	2220
	AGATATAGAG	TGACTATTTT	TAAAGCATGT	TAAAAATTTA	GGTTTTATTC	ATGTTTAAAG	2280
	TATGTATTAT	GTATGCATAA	TTTTTGCTGT	GTACTTGAAA	CTTAATCTTA	TCAAGAACTC	2340
35	TTTTTCATTG	ACTGAATGAT	TTCTTTTGCC	CCTAGGAGAA	AACTTAATAA	TTGTGCCTAA	2400
	AAACTATGGG	CGGATAGTAT	AAGACTATAC	TAGACAAAGT	GAATATTTGC	ATTTCCATTA	2460
	TCTATGAATT	AGTGGCTPAG	TTCTTTCTTA	GCTGCTTTAA	GGAGCCCCTC	ACTCCCCAGA	2520
	GTCAAAAAGG	AATGTAAAAA	CTTAGAGCTC	CCATTGTAAT	GTAAGGGGCA	AGAAATTTGT	2580
	GTCTCTCTGA	CTGAGCTGTC	CAGCACCAGC	CTTGTTTTAA	ATGTTTTCCT	GAGCTAGAAG	2640
40	AAATAGCTGA	TTATTGTATA	TGCAAAATTA	ATGCATTTTT	AAAAACTATT	CTTCTGTAAC	2700
	TTATCTACCT	GGTTATGATA	CTGTGGTCC	ATACACAAGT	AAAATAAGAT	TAGACAGAAG	2760
	CCAGTATACA	TTTTGCACATA	TTGATGTGAT	ACTGTAGCCA	GCCAGGACCT	TACTGATCTC	2820
	AGCATAATAA	TGCTCACTAA	TAATGAAGTC	TGCATAGTGA	CACCTCATCA	GACTGAAGAT	2880
	GAAGCAGGTT	ACGTGCTCCA	TTGGAAGGAG	TTTCTGATAG	TCTCTGCTG	TTTACCCCT	2940
45	TCCATTTTTT	AAAATAAGAA	ATTAGCAGCC	CTCTGCATAA	TGTAGCTGCC	TATATGCAGT	3000
	TTTTATCTGT	GCCCTAAAGC	CTCAGTGTCC	AGAGCTGTTG	GTCTATCAGT	GCTTATTGCA	3060
	CCCTCACCAT	GTGCCCTGGT	CCCTGCTGGG	TAGAGAACAC	AGAGGACAGG	GCATATCTCT	3120
	TGTCCTTAAG	GAGCTTGTGA	TCGTGTGACG	TAAAGCCCTC	TGGGATGTCT	GTGCCATGTG	3180
	ATTGACTTAC	AAGTGAAACT	GTCTTATAAT	ATGAAGGTCT	TTTTGTTTAC	TTCTAAACCC	3240
50	ACTTGGGTAG	TTACTATCCC	CAAACTGTGT	CTGTAAATAA	TATTATGGAA	GGGTTTCTAT	3300
	GTCACTCTAC	CTTAGAGAAA	GCCAGTGATT	CAATATCACA	AAAGGCATTG	ACGTATCTTT	3360
	GAAATGTTCA	CAGCAGCCTT	TTAAACAACA	CTGGGTGGTC	CTGTAGGCA	GAACATACTC	3420
	TCCTAAGTGG	TTGTAGGAAA	TTGCAAGGAA	AATAGAAGGT	CTGTCTTTCG	TCTCAAGGAG	3480
	GTTACCTTTA	ATAAAAGAAC	ACAAACCCAG	ATAGATATGT	AAACCAAAAT	ACTATGCCCC	3540
55	TTAATACTTT	TTAAGCAGCA	TTGTTAAATA	GTCTTACGCT	TTATACATTC	ACAGAATCTC	3600
	CCTGTTTTCC	TTGTATATAA	TGACTTTTGC	TGGCAGAACT	GAAATATAAA	CTGTAAAGGG	3660
	ATTTCGTGAG	TTGCTCCGAG	TATACAATAA	CCTCCAGGAC	ATAGCCAGAA	ATCTCCATTC	3720
	CACACATGAC	TGAGTTCCCTA	TCCCTGCAGT	GGTACTGGCT	CTTTTCTCCT	CTTTCTCTGC	3780
	CTCAGGGTTC	GTGCTACCCA	CTGATTCCCT	TTACCCTTAG	TAATAATTTT	GGATCATTTT	3840
60	CTTTCCCTTA	AAGGGGAACA	AAGCCTTTTT	TTTTTTTGAG	ACGGAGTGTG	GCTCTGTGAC	3900
	CCAAGCTGGA	GTGCAGTGGC	ACGATCTTGG	CTCACTCCAA	CCTCCACCTT	CCAGGTTCAA	3960
	GTGATTCTCC	TGCTCTCAGC	TCCCGAGTAG	CTGGGACTAC	GGGCACGCAC	CACCACGTCT	4020
	GGCTAATTTT	TGTATTTTTA	GTAGAGATGG	GGTTTCACCC	TATTTGGTCAG	GCTGGTCTTG	4080
	AATTCCTCAC	CTCAGGTCTAT	CCGCTGTGCT	CGGCTTCCCG	AAGTGCTGGG	ATTATAGGTG	4140
65	TGAGCCACCG	CACCCAGTTG	GGAACAAAGC	CTTTTAAACA	CACGTAAGGG	CCCTCAAACC	4200
	GTGGACCTC	TAAAGGAGAC	TTTGAAGCTT	TTTGAGGGCA	AACCTTACCT	TTGTGGTCCC	4260
	CAAAATGATG	CATTCTCTTT	TGAAATTTAT	TAGATACTGT	TATGTCCCCC	AAGGGTACAG	4320
	GAGGGGCATC	CTTAGCCCTA	TGGGAACACC	CAAACTAGGA	GGGGTTATTTG	ACAGGAAGGA	4380
	ATGAATCCAA	GTGAAGGCTT	TCTGCTCTTC	GTGTACAAAA	CCAGTTTCAG	AGTTAGCTTT	4440
	CTGGGGAGGT	GTGTGTTTTT	GAAAGGAATT	CAAGTGTGTC	AGGACAGATG	AGCTCAAGGT	4500
70	AAGGTAGCTT	TGGCAGCAGG	GCTGATACTA	TGAGGCTGAA	ACAACTCCCTG	TGATGAAGTA	4560
	GATCATGTCAG	TGACATACAA	AGACCAAGGA	TTATGTATAT	TTTTATATCT	CTGTGGTTTT	4620
	GAAACTTTAG	TACTTGAAT	TTTGGCCTTC	TGCACACTCT	TTTTGCTCTT	ACGAACATAA	4680
	TGGACTCTTA	AGAAATGAAA	GGGATGACAT	TTACCTATGT	GTGCTGCCTC	ATTCCTGGTG	4740
75	AAGCAACTGC	TACTTGTCTT	CTATGCCTCT	AAAATGATGC	TGTTTTCTCT	GCTAAAGGTA	4800
	AAAGAAAAAG	AAAAATAGT	TGGAATAATA	GACATGCAAC	TTGATGTGCT	TTTGAGTAAA	4860
	TTTATGCAGC	AGAAACTATA	CAATGAAGGA	AGAATTTCTAT	GGAAATTACA	AATCCAAAAC	4920
	TCATATGATG	TGCTTTCCCTA	GGGAGTAGAG	AAAGGCAGTG	AAATGGCAGT	TAGACCAACA	4980
	GAGGCTTGAA	GGATTCAAGT	ACAAGTAATA	TTTTGTATAA	AACATAGCAG	TTTAGGTTCC	5040
	CATAATCCTC	AAAAATAGTC	ACAAATATAA	CAAAGTTTCA	TGTTTTAGGG	TTTTTAAAAA	5100
80	ACGTGTTGTA	CCTAAGGCCA	TACTTACTCT	TCTATGCTAT	CACGTCAAGG	GGGTGATATG	5160

TATGTATTAT ATAAAAA AAACCCCTTAA TGCACCTGTTA TCTCCTAAAT ATTTAGTAAA 5220
 TTAATACATAT TTAATTTTTT TAAAGATTG TCTGTGTAGA CACTAAAAGT ATTACACAAA 5280
 ATCTGGACTG AAGGTGTCTT TTTTAAACAA AATTAAAGT ACTTTTATA TATGTTATGT 5340
 AGTATATCCT TTCTAAACTG CCTAGTTTGT ATATTCCAT ATTCCTTATT TGTGAAGTGT 5400
 ACCTGTTCTT GTCTCTTTT TCAATCATTT TCTGCACGCA TCCCCCTTATA TATGGTTATA 5460
 GAGATGACTG TAGCTTTTCG TGCTCCACTG CGAGGTTTGT GCTCAGAGCC GCTGCACCCC 5520
 AGCGAGGCCT GCTCCATGGA GTGCAGGACG AGCTACTGCT TTGGAGCGAG GGTTCCTGTC 5580
 TTTTGAGTTG ACCTGACTTC CTTCTTGAAA TGAATGTTAA AACTAAAATA AATTACATTG 5640
 CATTATTTT ATATTCTTGG TTGAAATAAA ATTTAATTGA CTTTG

SEQ ID NO:78 PDO3 Protein sequence:
 Protein Accession #: BAA82980

1 11 21 31 41 51
 VKSLLYQILD GIHYLHANWV LHRDLKPANI LVMGEGPERG RVKIADMGFA RLFNSPLKPL 60
 ADLDPVVVTF WYRAPELLLG ARHYTKAIDI WAIGCIFAEI LTSEPIHFHR QEDIKTSNPF 120
 HHDQLDRIFS VMGFPAKDW EDIRKMEYFP TLQKDFRRIT YANSSLIKYM EKHVKPDSK 180
 VFLLQLKLLT MDGCTKRTSE QALQDPYFQE DPLPTLDVFA GCQIPYPKRE FLNEDDPEEK 240
 GDKNQQQQN QHQQPTAPPQ QAAAPPQAPP PQONSTQING TAGGAGAGVG GTGAGLQHSQ 300
 DSSLNQVPPN KKPLRGPSGA NSGGPVMPSD YQHSSRLNY QSSVQSSQS QSTLYSSSS 360
 QQSSQYHPHSH QAHRY

SEQ ID NO:79 PDO5 DNA SEQUENCE

Nucleic Acid Accession #: XM_002922
 Coding sequence: 1-2190 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 ATGAATCCTT TCCAGAAAAA TGAGTCCAAG GAAACTCTTT TTTACCTGT CTCCATTGAA 60
 GAGGTACCAC CTCGACCACC TAGCCCTCCA AAGAAGCCAT CTCCGACAAT CTGTGGCTCC 120
 AACTATCCAC TGAGCATTGC CTTCAATTGT GTGAATGAAT TCTGCGAGCG CTTTTCCTAT 180
 TATGGAATGA AAGCTGTGCT GATCCTGTAT TCTCTGTATT TCTTGCACGT GAATGAAGAT 240
 ACCTCCACAT CTATATACCA TGCCCTTCAGC AGCCTCTGTT ATTTTACTCC CATCTGGGA 300
 GCAGCCATTG CTGACTCGTG GTTGGGAAAA TTTCAAGACAA TCATCTATCT CTCCTGGTG 360
 TATGTGCTTG GCCATGTGAT CAAGTCCCTG GGTGCCTTAC CAATACTGGG AGGACAAGTG 420
 GTACACACAG TCCTATCATT GATCGGCTG AGTCTAATAG CTTTGGGGAC AGGAGGCATC 480
 AAACCTGTG TGCGAGCTTT TGGTGGAGAC CAGTTTGAAG AAAAACATGC AGAGGAACGG 540
 ACTAGATACT TCTCAGTCTT CTACCTGTCC ATCAATGCAG GGAGCTTGAT TTCTACATTT 600
 ATCACACCCA TGCTGAGAGG AGATGTGCAA TGTTTTGGAG AAGACTGCTA TGCATTGGCT 660
 TTTGGAGTTC CAGGACTGCT CATGGTAATT GCACCTGTGT TGTTTGCAAT GGAAGCAAAA 720
 ATATACAAATA AACCAACCCC TGAAGGAAC ATAGTGGCTC AAGTTTTCAA ATGTATCTGG 780
 TTTGCTATTT CCAATCGTTT CAAGAACCGT TCTGGAGACA TTCCAAGCG ACAGCACTGG 840
 CTAGACTGGG CAGCTGAGAA ATATCCAAAG CAGCTCATTA TGGATGTAAA GGCAGTACC 900
 AGGGTACTAT TCCTTTTATAT CCCATTGCCC ATGTTCTGGG CTCCTTTTGA TCAGCAGGCT 960
 TCACGATGGA CTTTGCACAG CATCAGGATG AATAGGAATT TGGGGTTTGT TGTGCTTCAG 1020
 CCGGACCAGA TGCAGGTTCT AAATCCCTTT CTGGTCTTFA TCTTCATCCC GTTGTGTTGAC 1080
 TTTGTCATTT ATCGTCTGTT GGAATTAAT TCTCATCACT TAGGAAAAATG 1140
 GCTGTGTTGA TGATCTTAGC GTGCCTGGCA TTTGCAGTTG CGGCAGCTGT AGAGATAAAA 1200
 ATAAATGAAA TGGCCCCAGC CCAGTCAGGT CCCAGGAGG TTTTCCTACA AGTCTTGAAT 1260
 CTGGCAGATG ATGAGGTGAA GGTGACAGTG GTGGGAAATG AAAACAATTC TCTGTTGATA 1320
 GAGTCCATCA AATCCTTTCA GAAAACACCA CACTATTCCA AACTGCACCT GAAAACAAAA 1380
 AGCCAGGATT TTCACCTCCA CCTGAAATAT CACAATTTGT CTCCTACAC TGAGCATCTT 1440
 GTGCAGGAGA AGAACTGGTA CAGTCTGTTC ATTCGTGAAG ATGGGAACAG TATCTCCAGC 1500
 ATGATGGTAA AGGATACAGA AAGCAAAACA ACCAATGGGA TGACAACCGT GAGGTTTGT 1560
 AACACTTTGC ATAAAGATGT CAACATCTCC CTGAGTACAG ATACCTCTCT CAATGTTGGT 1620
 GAAGACTATG GTGTGTCTGC TTATAGAAGT GTGCAAGAG GAGAATACCC TGCACTGCAC 1680
 TGTAGAACAG AAGATAAGAA CTTTCTCTG AATTTTGGGT TTCTAGACTT TGGTGCAGCA 1740
 TATCTGTTTG TTATTACTAA TAACACCAAT CAGGCTCTTC AGGCCTGGAA GATTGAAGAC 1800
 ATTCCAGCCA ACAAAATGTC CATTGCGTGG CAGCTACCAC AATATGCCCT GGTTCAGACT 1860
 GGGGAGGTGA TGTCTCTGT CACAGTCTT GAGTTTCTT ATTCTCAGGC TCCCTCTAGC 1920
 ATGAAATCTG TGCTCCAGGC AGCTTGGCTA TTGACAATTG CAGTTGGGAA TATCATCGTG 1980
 CTGTGTGTGG CACAGTTCAG TGGCCTGGTA CAGTGGGCGG AATTCATTTT GTTTTCTGTC 2040
 CTCCTGCTGG TGATCTGCTT GATCTTCTCC ATCATGGGCT ACTACTATGT TCCTGTAAAG 2100
 ACAGAGGATA TGCGGGGTCC AGCAGATAAG CACATTCCTC ACATCCAGGG GAACATGATC 2160
 AACTAGAGA CCAAGAAGAC AAAACTCTGA

SEQ ID NO:80 PDO5 Protein sequence:
 Protein Accession #: XP_002922

1 11 21 31 41 51
 MNPFOKNESE ETLFSEVSI EPPRPPSP KPSPTICGS NYPLSIAFIV VNEFCERFSY 60
 YGMKAVLILY FLYFLHWNQ TSTSIYHAFS SLCYPTPILG AAIADSWLKG FKTIILSLV 120
 YVLGHVILSL GALPILGGQV VHTVLSLIGL SLIALGTGGI KPCVAAFGGD FEEKHAEER 180
 TRYFVSFVYLS INAGSLISTF ITPMLRGDVQ CFGEDCYALA FGVPGLLMVI ALVVFAMGSK 240
 IYNKPPPEGN IVAQVEKCIW FAISNRFKNR SGDIFKQHW LDWAAEKYPK QLIMDVKALT 300
 RVLFYIPLP MFWALLDQQG SRWTLQAI RMNRNLGFFVLQ PDQMQLVNF LVLIFIPFLD 360
 FVIYRLVSK GINFSRLRM AVGMILACIA PAVAAVEIK INEMAPAQSG PQEVFLQVLN 420
 LADDEVKVTY VGNENNSLLI ESIKSFQKTP HYSKHLHLTK SQDFHFLHY HNLISLYTEHS 480

VQEKWYSLV IREDGNSISS MMVKDTESKT TNGMTTVRFV NTLHKDVNIS LSTDTSLNVG 540
EDYGVSAVRT VQRGEYPAVH CRTEDKNFSL NLGLLDFGAA YLFVITNNNTN QGLQAWKIED 600
IPANKMSIAW QLPQYALVTA GEVMFSVTGL EFSYSQAPSS MKSVLQAAML LTIAGVNIIV 660
LVVAQFSLV QWAEFILFSC LLLVICLIFS IMGYYYVPVK TEDMRGPADK HIPHIQGNMI 720
KLETKKTKL

SEQ ID NO:81 PDO6 DNA SEQUENCE

Nucleic Acid Accession #: NM_020448

Coding sequence: 1-1221 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
| | | | |
ATGGACGGAT CCCACAGCGC AGCCCTGAAG CTGCAGCAGC TGCCCTCCAC AAGTAGCTCC 60
AGCGCCGTAA GCGAGGCGCT CTCTCTCTAC AAGGAAAACC TGATTGGCGC CCTCTGGCGG 120
ATCTTCGGGC ACCTCGTGGT CAGCATTGCA CTTAACCTCC AGAAGTACTG CCACATCCGC 180
CTGGCAGGCT CCAAGGATCC CCGGCGCTAT TCAAGACCA AGACATGGTG GCTGGGCTG 240
TTCTGTATGC TTCTGGGCGA GCTGGGTGTG TTGCGCTCCT ACGCCTTCGC GCCGCTGTCA 300
CTCATCGTGC CCCTCAGCGC AGTTTCTGTG ATAGCTAGTG CCATCATAGG AATCATATTC 360
ATCAAGGAAA AGTGGAAACC GAAAGACTTT CTGAGGCGCT ACGTCTGTG CTTTGTGGC 420
TGCGGTTTGG CTGTCGTGGG TACCTACCTG CTGGTGACAT TCGCACCCAA CAGTCACGAG 480
AAGATGACAG GCGAGAATGT CACCAGGCAC CTCGTGAGCT GGCCTTTCCT TTTGTACATG 540
CTGTGTGGAG TCATTCGTGT CTGCTTGCTG CTCTACTTCT ACAAGGAGAA GAACGCCAAC 600
AACATTGTGC TGATTCTTCT CTTGGTGGCG TTACTTGGCT CCATGACAGT GGTGACAGTC 660
AAGCGCGTGG CTGGGATGCT TGTCTTGTC ATTCAAGGGA ACCTGCAGCT TGAATACCCC 720
ATCTTCTACG TGATGTTCTG GTGCATGGTG GCAACCGCCG TCTATCAGGC TCGGTTTTTG 780
AGTCAAGCCT CACAGATGTA CGACTCCTCT TTGATTGCCA GTGTGGGCTA CATTCTGTCC 840
ACAACCATTT CTATCACAGC AGGTGCAATA TTTTACCTGG ACTTCATCGG GGAGGACGTG 900
CTGCACATCT GCATGTTTGC ACTGGGGTGC CTCATTGCAT TCTTGGGCGT CTTCTTAATC 960
ACGCGTAACA GGAAGAAGCC CATTCATTTC GAGCCCTATA TTTCCATGGA TGCCATGCCA 1020
GGTATGCAGA ACATGCACGA TAAAGGGATG ACTGTCCAGC CTGAACCTAA AGCTTCTTTT 1080
TCCATATGGG CTCTGGAAAA CAATGACAA ACATTCTGAGA TCTACGCTCC TGCCACCCCTG 1140
CCAGTCATGC AAGAAGAGCA CGGCTCCAGA AGTGCCCTCTG GGGTCCCTTA CCGATCCTTA 1200
GAGCACACCA AGAAGGAATG A

SEQ ID NO:82 PDO6 Protein sequence

Protein Accession #: NP_065181

1 11 21 31 41 51
| | | | |
MDGSHSAALK LQQLPPTSSS SAVSEASFSY KENLIGALLA IFGHLVVSIA LNLQKYCHIR 60
LAGSKDPRAY FKTKTWLGLV FLMLLGLGV FASYAFAPLS LIVPLSAVS IASAIIGIIF 120
IKEKWKPKDF LRRYVLFSVG CGLAVVGTYL LVTFAFNSHE KMTGENVTRH LVSWPFLLYM 180
LVEIILFCLL LYFYKEKNAN NIVVILLLVA LLGSMTVVTV KAVAGMLVLS IQGNLQLDYP 240
IFYVMFVCMV ATAVYQAAFL SQASQMYDSS LIASVGYILS TTIAITAGAI FYLDFIGEDV 300
LHICMFALGC LIAFLGVFLI TRNRKKPIPF EPYISMAMP GMQNMHDKGM TVQPELKASF 360
SYGALENNND ISEIYAPATL FVMQEEHGSR SASGVYPRVL EHTKKE

SEQ ID NO:83 PDO8 DNA SEQUENCE

Nucleic Acid Accession #: NM_032712

Coding sequence: 555-908 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
| | | | |
CACTCATTAA GAACAGAGGA GGCTGCCTGT TACTCTCTGGT GTTGCATCCC TCCAGACACT 60
CTGCTGTTC CTGCTTAGGC GTGGCTGCAG CCATGGCTAG GAAAGCGCTG CCACCCACCC 120
ACCTGGGCCA GAGCTGGTTC TGCTCTGCT GCAGGGACAC TGAGCTGGCT ATCTCGGCGC 180
TTCGGGCAAG AACTGCAACA GGCTCTCCTG GGTCTGCAG GTGTACAGCC GGGCCCTGTC 240
CTTGTGCCCT AGCTCTCGAG AGCTGCTGCT GCGGGGTGAC CTGATCCAAC CTGATAAGGT 300
GCCATCTTCA GCTACCACTG CAAGGCCCTG AGGGCAACAG CAGCAGCGCA CTGCCCCACC 360
GGCTGCTGAT GGCCTGGTGC CAGCTGGGAG TCCTCCCGGC ACTTCGAGGC CACTGAGCCA 420
CCCTTCCAGC CCCAGCCAC CATGGACAGG GGTATCCAGC TTCTCTCTCA ACCTCGTCTT 480
CTGCCCCTGA GCCAGTGACG CCAAGGACA TGCTGTGTAC CCAGGTCTCT TACCAGCACT 540
AGCTGGTCAA GGGCATGACA GTGCTGGAGG CCGTCTTGGA GATCCAGGCC ATCACTGGCA 600
CGAGGCTGCT CTCCATGGTG CCAGGGCCCG CCAGGCCACC AGGCTCATGC TGGGACCCAA 660
CCCACTGCAC AAGGACTTGG CTGCTGAGCC ACACACCCAG GAGAAGGTGG ATAAGTGGGC 720
TACCAAGGGC TTCTGCAAG CTAGGGGAGG AGCCACCCCC GCTTCCCTAT TGTGACCAGG 780
CCTATGGGGA GGAGCTGTCC ATACGCCACC GTGAGACCTG GGCCTGGCTC TCAAGGACAG 840
ACACCGGCTG GCTCTGTGCT CCAGGGGTGA AGCAGGCCAG AATCCTGGGG GAGCTGTCTC 900
TGTTTGTAGC TGCAATTCAGG AAGTGCGGGA CATGGTAGGG GAGGCAAAAA GCCTTGGGCA 960
CTACCCCTCC TGTTGAGCTG TTGGGTGTCC GTCGAGTAG CCACACCCCTG ACACCATGTT 1020
CAAGGGTACC GGAAGAGAAG GGTGTCTGCC CCCAACCTCC CCTGTGGGTG TCACTGGCCA 1080
GATGTATGA GGAAGCAGG CCTGTGTAGT GGACACTGAC CATGAGTCCC TGGGGGGAGT 1140
GATCCCCAG GCATCGTGTG CCATGTTGCA CTTCTGCCCA GGCAGCAGGG TGGGTGGGTA 1200
CCATGGGTGC CCACCCCTCC ACCACATGGG GCCCCAAAGC ACTGCAGGCC AAGCAGGGCA 1260
ACCCACACCC CTTGACATAA AAGCATCTTG AAGCTTTTAA AAAAAAAAAA AAAAAA

SEQ ID NO:84 PDO8 Protein sequence

Protein Accession #: NP_116101

1 11 21 31 41 51

MTVLEAVLEI QAITGSRLLS MVEGPARPPG SCWDPTQCTR TWLLSHTPRR RWISGLPRAS 60
CRLGEEPPPL PYCDQAYGEE LSIRHRETWA WLSRTDTAWP GAPGVKQARI LGLELLLV

5

SEQ ID NO:85 PDT1 DNA SEQUENCE

Nucleic Acid Accession #: NM_000693

Coding sequence: 53-1591 (underlined sequences correspond to start and stop codons)

10 1 11 21 31 41 51
AGCCGGTGGC CCGCAGACTA GGGCGCCTCG GGCCAGGGAG CGCGGAGGAG CCATGGCCAC 60
CGCTAACGGG GCCGTGGAAA ACGGGCAGCC GGACGGGAAG CCGCCGGCCC TGCCGCGCCC 120
CATCCGCAAC CTGGAGGTCA AGTTACCAA GATATTTATC AACAAATGAAT GGCACGAATC 180
CAAGAGTGGG AAAAAGTTTG CTACATGTAA CCTTCAACT CGGGAGCAAA TATGTGAAGT 240
15 GGAAGAAGGA GATAAGCCCG ACGTGGACAA GGCTGTGGAG GCTGCACAGG TTGCTTCCA 300
GAGGGGCTCG CCATGGCCCG GGCTGGATGC CCTGAGTCGT GGGCGGCTGC TGCAACAGCT 360
GGCTGACCTG GTGGAGAGGG ACCGCGCCAC CTTGGCCGCC CTGGAGACGA TGGATACAGG 420
GAAGCCATTT TTTCATGCTT TTTTCATCGA CCTGGAGGGC TGTATTAGAA CCCTCAGATA 480
CTTTGCAGGG TGGGCAGACA AAATCCAGGG CAAGACCATC CCCACAGATG ACAACGTCGT 540
20 ATGCTTCACC AGGCATGAGC CCATTTGGTGT CTGTGGGGCC ATCACTCCAT GGAACCTCCC 600
CCTGCTGATG CTGTGTGGGA AGCTGGCACC CGCCCTCTGC TGTGGGAACA CCATGGTCTC 660
GAAGCCTGCG GAGCAGACAC CTCTCACCGC CCTTTATCTC GGCTCTCTGA TCAAAGAGGC 720
CGGGTTCCCT CCAGGAGTGG TGAACATTGT GCCAGGATTC GGGCCACAG TGGGAGCAGC 780
AATTTCTTCT CACCCTCAGA TCAACAAGAT CGCCTTCACC GGCTCCACAG AGGTTGGAAA 840
25 ACTGGTTAAA GAAGCTGCGT CCCGGAGCAA TCGAAGCGGG GTGACGCTGG AGCTGGGGGG 900
GAAGAACCCC TGCATCGTGT GTGCGGACGC TGACTTGGAC TTGGCAGTGG AGTGTGCCCA 960
TCAGGGAGTG TTCTTCAACC AAGGCCAGTG TTGCACGGCA GCCTCCAGGG TGTTCGTGGA 1020
GGAGCAGGTC TACTCTGAGT TTGTACGGCG GAGCGTGGAG TATGCCAAGA AACGGCCCGT 1080
GGGAGACCCC TTGCATGTCA AAACAGAAAC GGGGCCCTCAG ATTGATCAAA AGCAGTTCGA 1140
30 CAAAATCTTA GAGCTGTATG AGAGTGGGAA GAAGGAAGGG GCCAAGCTGG AATGCGGGGG 1200
CTCAGCCATG GAAGACAAGG GGCTCTTCAT CAAACCCACT GTCTTCTCAG AAGTCACAGA 1260
CAACATGCGG ATTGCCAAAG AGGAGATTTT CGGGCCAGTG CAACCAATAC TGAAGTTCAA 1320
AAGTATCGAA GAAGTGATAA AAAGAGCGAA TAGCACCGAC TATGGACTCA CAGCAGCCGT 1380
GTTCAACAAA AATCTCGACA AAGCCCTGAA GTTGGCTTCT GCCTTAGAGT CTGGAACGGT 1440
35 CTGGATCAAC TGCTACAACG CCCCTCTATG ACAGGCTCCA TTTGGTGGCT TTAATAATGTC 1500
AGGAAATGGC AGAGAACTAG GTGAATACGC TTTGGCCGAA TACACAGAAG TGAAACTGT 1560
CACCATCAAA CTTGCGGACA AGAACCCCTG AAGGAAAGGC GGGGCTCCTT CCTCAACAT 1620
CGGACGGCGG AATGTGGCAG ATGAAATGTG CTGGAGGAAA AAAATGACAT TTCTGACCTT 1680
40 CCCGGGACAC ATTCTCTTGG AGGCTTTACA TCTACTGGAG TTGAATGATT GCTGTTTTC 1740
TCTCACTCTC CTGTTTATTC ACCAGACTGG GGATGCTTAT AGGTTGTCTG TGAATCGCA 1800
GTCCTGCCCG GGGAGGGAGC TGTGGCCAT TTTCTGTGTT CCCTTTAAAC CAGATCTCTG 1860
AGACAGTGA ATACTCAGGG CGTTGTTAAC AGGGAGTGGT ATTTGAAGTG TCCAGCAGTT 1920
GCTTGAATG CTTTGCAGAA TCTGACTCCA GTAAGAATGT GGGAAAACCC CTTGTGTGTT 1980
45 CTGCAAGCAG GGCTCTTGCA CCAGCGGTCT CCTCAGGGTG GACCTGCTTA CAGAGCAAGC 2040
CAGCGCTCTT TCCGAGGTGA AGGTGGGACC ATTCCTTGGG AAAGGATTC AAGTAAAGTT 2100
TTTGGGTTT TGTGTTTGT TTTCTTGT TTAATAAAG GATTTCACAG TGAGAAAGTT 2160
TTGGTTAGTG CATACCGTGG AAGGCGCCA GGGTCTTGT GGATTGCATG TTGACATTGA 2220
CCGTGAGATT CGGCTTCAAA CCAATACATG CTTTGGAAAT TGACAGAATC AATAGCCAG 2280
50 AGAGCTTAGT CAAAGACCAT ATCAGCGTCT ACCTTAACCA AGGCACTTTC TTAAGCAGAA 2340
AATATGTTG AGGTTACCTT TGCTGTCTAA GATCCAATCT TCTAACGCCA CAACAGCATA 2400
GCAAAATCTA GGATATTTCA CCTCTCAT TGAACAAATC GAGCTGTAAT TCACTTTAAC 2460
AAATTACGCA TTTCTATCAC GTTCACTAAC AGCTTATGAT AAGTCTGTGT AGTCTTCCTT 2520
TTCTCCAGTT CTGTTACCCA ATTTAGATTA GTAAAGCGTA CACAACCTGA AAGACTGCTG 2580
55 TAATAACACA GCCCTGTTAT TTTTAAAGTCC TATTTTGATA TTAATTTCTG ATTAGTTAGT 2640
AAATAACACC TGGATCTTAT GGAGGACCTC GGTCTTCTAT CAAGTGGCCT GAGTATTTCA 2700
CTGGCAGGTT GTGAATTTT CTTTCTCTCT TTGGGAATCC AAATGATGAT GTGCAATTTT 2760
ATGTTTAAAC TTGGGAAACT GAAAGTGTTC CCATATAGCT TCAAAAACAA AAACAAATGT 2820
GTTATCCGAC GGATACTTTT ATGGTTACTA ACTAGTACTT TCCTAATTTG GAAAGTAGTG 2880
60 CTTAAGTTTG CAAATTAAGT TGGGGAGGGC AATAATAAAA TGAGGGCCCG TAACAGAACC 2940
AGTGTGTGTA TAACGAAAC CATGTATAAA ATGGGCCAT CACCCTTGTC AGAGATATAA 3000
ATTACCACAT TTGGCTTCCC TTCATCAGCT AACACTTATC ACTTATACTA CCAATAACTT 3060
GTTAAATCAG GATTTGGCTT CATACTGTA ATTTTCAGTA TTTTATCTCA AGTAGATATA 3120
GACACTAACC TTGATAGTGA TACGTTAGAG GGTTCCTATT CTTCATATGT ACGATAATGT 3180
65 CTTTAAATAG AAATGCTACA TTATTTATAA TTGGTAGAGT TATTTGATCT TTTTATAGTT 3240
GTAAGTACAC AGAGGTGGTA TATTTAACT TCTGTAATAT ACTGTATTTA GAAATGGAAA 3300
TATATATAGT GTTAGGTTTC ACTTCTTTTA AGGTTTACCC CTGTGGTGTG GTTTAAAAAT 3360
CTATAGGCTT GGGAAATCCG ATCCTAGCTG CAGATCGCAT CCCCAATGCG GAGAATGATA 3420
AAATAAAATT GGATATTGGA GA

70 SEQ ID NO:86 PDT1 PROTEIN SEQUENCE

Protein Accession #: NP_000684

75 1 11 21 31 41 51
MATANGAVEN GQPDGKPPAL PRPIRNLEVK FTKIFINNEW HESKSGKKFA TCNPSTREQI 60
CEVEEGDKPD VDKAVEAAQV AFQGSPPWRR LDALSRGRL HQLADLVERD RATLAALETM 120
DTGKPFLLHF FIDLEGCIRT LRYFAGWADK IQGKTIPTDD NVVCFTRHEP IGVCGAITPW 180
NFPLMLVWK LAPALCCGNT MVLKPAEQTP LTALYLGSLI KEAGFFPGVV NIVPGFGPTV 240
GAAISSHPQI NKIAFTGSTV VGKLVKERAS RSNLKRVTLE LGGKNPCIVC ADADLDLAVE 300

CAHQGVFFNQ GQCCTAASRV FVEEQVYSEF VRRSVEYAKK RFPVGDPPFDVK TEQGPQIDQK 360
 QFDKILELIE SGKKEGAKLE CGGSAMEDKG LFIKPTVFSE VTDNMRIAKE EIFGFVQPIL 420
 KFKSIEEVIK RANSTDYGLT AAVFTKNLDK ALKLASALEK GTVWVWVYNA LYAQAPFGGF 480
 KMSGNGRELK EYALAEYTEV KTVTTIKLGDK NP

5

SEQ ID NO:87 PDV3 DNA SEQUENCE

Nucleic Acid Accession #: NM_032642

Coding sequence: 184-1263 (underlined sequences correspond to start and stop codons)

10 1 11 21 31 41 51
 | | | | | |
 GACCATTAGC AGGCACCCAG GCCTGTCTTT GGCTCGGAAA CGGTGGCCCC CAATGTAGCC 60
 TAGTTTGAAC CTAGGAACAG CAGGACCAGA GAGATTCCAC TGGAGCCTGA TGGACGGGTG 120
 ACAGAGGGAA CCCTACTCTG GAAACTGTCA GTCCCAGGGC ACTGGGGAGG GCTGAGGCCG 180
 15 ACCATGCCCA GCCTGCTGCT GCTGTTCACG GCTGCTCTGC TGTCCAGCTG GGCTCAGCTT 240
 CTGACAGACG CCAACTCCTG GTGGTCATTA GCTTTGAACC CGGTGCAGAG ACCCGAGATG 300
 TTTATCATCG GTGCCCAGCC CGTGTGCAGT CAGCTTCCCG GGCTCTCCCC TGGCCAGAGG 360
 AAGCTGTGCC AATTGTACCA GGAGCACATG GCCTACATAG GGGAGGGAGC CAAGACTGGC 420
 ATCAAGGAAT CCAGACACCA GTTCCGGCAG CGCGGTGGA ATTGCAGCAC AGCGGACAAAC 480
 20 GCATCTGTCT TTTGGGAGAGT CATGCAGATA GGCAGCCGAG AGACCGCCTT CACCCACGCG 540
 GTGAGCGCCG CGGGCGTGGT CAACGCCATC AGCCGGGCCCT GCCCGAGAGG CGAGCTCTCC 600
 ACCTGCGGCT GCAGCCGGAC GCGCGCGCCC AAGGACCTGC CCCGGGACTG GCTGTGGGGC 660
 GGCTGTGGGG ACAACGTGGA GTACGGCTAC CGCTTCGCCA AGGAGTTTGT GGATGCCCGG 720
 GAGCGAGAGA AGAATTTTGC CAAAGGATCA GAGGAGCAGG GCCGGGTGCT CATGAACCTG 780
 25 CAAAACAAAG AGGCCGTGCG CAGGGCTGTG TATAAGATGG CAGACGTAGC CTGCAAAATG 840
 CACGGCGTCT CGGGTCTCTG CAGCCTCAAG ACCTGCTGGC TGCAGCTGGC CGAGTTCCGC 900
 AAGGTGCGGG ACCGCTGTGA GGAGAAAGTAC GACAGCGCGG CCGCCATGCG CGTCAACCGC 960
 AAGGGCCGGC TGGAGCTGGT CAACAGCCGC TTCACCCAGC CCACCCCGGA GGACCTGGTC 1020
 TAGTGTGACC CCAGCCCCGA CTACTGCCCTG CGCAACGAGA GCACGGGCTC CCTGGGCACG 1080
 30 CAGGGCCGCC TCTGCAACAA GACCTCGGAG GGCATGGATG GCTGTGAGCT CATGTGCTGC 1140
 GGGCGTGGCT ACAACCAAGT CAAGAGCGTG CAGGTGGAGC GCTGCCACTG CAAGTTCAC 1200
 TGGTGTCTCT TCCTCAGGTG TAAGAAGTGC ACGGAGATCG TGGACCACTA CATCTGTAAA 1260
 TAGCCCGGAG GGCCTGCTCC CGGCCCCCCC TGCACCTGCG CTCACAAAGG TCTATATTAT 1320
 35 ATAAATCTAT ATAAATCTAT TTTATATTG TATAAGTAAA TGGGTGGGTG CTATACAAATG 1380
 GAAAGATGAA AATGGAAGG AAGAGCTTAT TTAAGAGACG CTGGAGATCT CTGAGGAGTG 1440
 GACTTTGCTG GTTCTCTCTC CTTGGTGGGT GGGAGACAGG GCTTTTCTC TCCCTCTGGC 1500
 GAGGACTCTC AGGATGTAGG GACTTGGAAA TATTTACTGT CTGTCCACCA CGGCCTGGAG 1560
 GAGGGAGGTT GTGGTTGGAT GGAGGAGATG ATCTTGTCTG GAACTCTAGA GTCTTTGTG 1620
 40 GTTAGAGGAC TGCCCTGTAT CCTGGCCACT AGGCCAAGAG GCCCTATGAA GGTGGCGGGA 1680
 ACTCAGCTTC AACCTCGATG TCTTCAAGGT CTGTCCAGA ATGTAGATGG GTTCCGTAAG 1740
 AGGCTTGGTG CTCTCTTACT CTTTCATCCA CGTGCACTTG TGGCGCATCT GCAGTTTACA 1800
 GGAACGGCTC CTTCCTTAAA ATGAGAAGTC CAAGGTTCATC TCTGGCCAG TGACACAGA 1860
 GAGATCTGCA CCTCCCGGAG TPCAGGCTG CTTTCCAGC GAGAATTCTT CATCTCCAC 1920
 45 GGTTCACCTAG CTCTTACCTG AAGAGGAAAG GGGGCCATTG GACCTGACAT GTCAGGAAAG 1980
 CCCTAAACTG AATGTTTGGC CCTGGGCTGC AGAAGCCAGG GTGCATGACC AGGCTTGGTG 2040
 GACGTATATC TGTCTTCCCC CACCCCGGGG TTGAGCTGCT GTGTCTACTC 2100
 CTCCACCGAG GGAGGCTCA CAAACACAG GACGCTGCAA CGGCTCAGGC TGGCGGGCCC 2160
 GCGGTGCTCA TCATCTCTGC CCCAGGTGTA CGGTTTCTCT CTGACATTAA ATGCCCTTCA 2220
 50 TGGAAAAAAA AAAAAGAAAA AAAAAAAAAA AA

SEQ ID NO:88 PDV3 Protein sequence

Protein Accession #: NP_116031

55 1 11 21 31 41 51
 | | | | | |
 MPSSLLLFTH ALLSSWAQLL TDANSWWSLA LNPVORPEMF IIGAQPVCQSQ LPGLSPGQRK 60
 LCQLYQEHMA YIGEGAKTGI KECQHQRQR RNWCSTADNA SVFGRVMQIG SRETAFTHAV 120
 SAAGVVNAIS RACREGELST CGCSRTARPK DLPRDWLWGG CGDNVEYGYR FAKEFVDARE 180
 60 REKNFAKGSE EQGRVLMNLQ NNEAGRRVY KMADVACKCH GVSQSCSLKT CWLQLAEFRK 240
 VGDRLEKEKYD SAAAMRVTRK GRLELVNSRF TQFTPEDLVY VDPSPDYCLR NESTGSLGTQ 300
 GRLCNKTSSE MDGCELMCCG RGYNQFQSVQ VERCHCKFWH CCFVRCKKCT EIVDQYICK-

SEQ ID NO:89 PDT9 DNA SEQUENCE

Nucleic Acid Accession #: NM_033280

Coding sequence: 58-636 (underlined sequences correspond to start and stop codons)

70 1 11 21 31 41 51
 | | | | | |
 GGCAGCCGTC TGTGCCACCC AGAGCCGGCG GGCCGCTAGG TCCCCGGAGA CCCTGCTATG 60
 GTGCGTGGCG GCGCCGTGGG GGCTCATCTC CCCCGCTCCG GCTTGGATAT CTTGCGGGAC 120
 CTGAAGAAGA TGAAACAAGC CCAGCTCTAT TACCAGGTTT TAAACTTCGC CATGATCGTG 180
 TCTTCTGCAC TCATGATATG GAAAGGCTTG ATCGTGCTCA CAGGCAGTGA GAGCCCCATC 240
 GTGGTGGTGC TGAGTGGCAG TATGGAGCCG GCCTTTCACA GAGGAGACCT CCTGTTCTCT 300
 75 ACAAATTTCC GGAAGACCC AATCAGAGCT GGTGAAATAG TTGTTTTTAA AGTTGAAGGA 360
 CGAGACATTC CAATAGTTCA CAGAGTAATC AAAGTTTATG AAAAAGATAA TGGAGACATC 420
 AAATTTCTGA CTAAAGGAGA TAATAATGAA GTTGTATGATA GAGGCTTGTA CAAAGAAGGC 480
 CAGAAGCTGC TGGAAAAGAA GGACGTGGTG GGAAGAGCAA GAGGGTTTTT ACCATATGTT 540
 GGTATGGTCA CCATAATAAT GAATGACTAT CCAAAATTC AATATGCTCT TTTGGCTGTA 600
 80 ATGGGTGCAT ATGTGTTACT AAAACGTGAA TCCTAAATG AGAAGCAGTT CCTGGGACCA 660
 GATTGAAATG AATTCTGTTT AAAAAGAGAA AAACATAATAT ATTTGAGATG TTCCATTTTC 720

TGTATAAAAG GGAACAGTGT GGAGATGTTT TTGTCTTGTC CAAATAAAAG ATTCACCAGT 780
 AAAAAAAAAA AAAA

SEQ ID NO:90 PDT9 Protein sequence

Protein Accession #: NP_150596

1 11 21 31 41 51
 | | | | |
 MVRAGAVGAH LPASGLDIFG DLKMKMKRQL YYQVLNFAMI VSSALMIWKV LIVLTGSESP 60
 10 IVVVLSGSME PAFHRGDLFF LTNFREDFIR AGEIVVFKEV GRDIPIVHRV IKVHEKDNGD 120
 IKFLTCKDNN EVDDRGLYKE GQNWLEKKDV VGRARGFLPY VGMVTIIMND YPKFKYALLA 180
 VMGAYVLLKR ES

SEQ ID NO:91 PDV5 DNA SEQUENCE

Nucleic Acid Accession #: NM_016590

Coding sequence: 691-975 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 GATTACTCAC ACAGTCTTGA AGATGCAATG TCAGCTATTT AGGACAGAAA CATCCAAGGC 60
 CGTGTACGAA CTCAATTACG ACTACATATG CATTAAGGCA GGAAGTGGCA GGCCCTCAGGG 120
 TACGCCAACT ATAGGACTCG TGCTTCTCGT ACGCTGGGCT ATAATCTATG AAAGTGAAGT 180
 CCAGAGCCAG CCAATCACTT AGCTCCTCAT AACAAGTCTA ACTGGCTCTG GAAAGCTGAA 240
 AGGGCTGCAC TGAACAAACA CAGATGAGAT ATTCTACACA TTAATCTACT TATCTGGAAT 300
 25 CACTTTGCGT CTAAGAGCCA GAGAAAAATC ACAGCTTCCT TGTCCGAGGG GAAAAGGACA 360
 GGTGATCTGG GGAAGACGCA GGTACACCTG GAGCAAGGTC TCTTCCCGGC TTGGCAATCT 420
 CAGCTGTGCC GGCCTACGCG GACCCGAGCC GTCCCAGAAA CCAAGGGCA GGCACGGCAG 480
 CAAACGCCTG AGTGCTGCTG CCTTCGGTGA CTATATGAGA ATGGAAGCTT CTAAGGAAGC 540
 CAGGTGTGTA GAATGTGTAC CCCCTTTACT CAGAGATAAC ATAGATTATC CAGGCTGAGA 600
 30 TGGAAAACAA GCCCTTTATT GAATTTTCAA CACAGACTCC CTGCTTCTCA TCTCCTTAAT 660
 AAAATTTTAT TAAATCCCC TTGAACCTCC ATGTTCAAAAT CTCCATTGTG TGACAGACAA 720
 AGCCAACAAT ACTCTAAACT GAGGCCCTGCA AGTCATTTC A TTTGTATTTT TGTCCAGAAA 780
 TTTCCCATAG GAAGACTTCA CCTCCTACAA CTCCGAAGAA AACCCTTACT TCCCAAGACC 840
 35 GTCAACAGCA ACCATCCGCA GTCAATCAAG TGGAAGCTTT CACAGCTTTT GTACATTCTC 900
 TGTGTCAATA TACAACCTAG TTACAGACTG TCCCTGGCT CCTGACCCT TACAAACACT 960
 AAAAGTTTGT TTTGACTCAA CTTCAAGCTG CTCATCTGTT AGTAAGTGTAT GTTCACTCCA 1020
 GAACACATTC ATGATGAGAA CTTTCTAAAA GACCAGCACT GCTCTTCCCC TCCTATAATC 1080
 ATAATAATCA TGATAAAGCT AAACATGTTA CTGGGACTCG ACATTTTCTT GGGGATTGAA 1140
 40 ATCTTTAGTC CTTGGAGCTG TCACATAGCA GGGGCAACCT CACACTGAAA CAAAGGAAAT 1200
 GATGTCCCAT TATTATCCAC CCTGAGCCAC CATAATATGC TGTTTACATT TATTTTCTTC 1260
 AGCCTGTGCA AAACAAAGCA ATGGAAAAGG AAACATAAAA ATATACATAC TAGTACCATT 1320
 ATCTTCTTTT GCCTAAATTT ACTAATGCAC CACGTCAAGT TGCTTCTTTC AGGCATCATT 1380
 CTCAATTCAT CAGGACTTGT ATTAGCAGGT TCTGGCTAGA GAGACTATCT CCTGTATCA 1440
 45 CGATCAATTA ATGTTTCTGT GTGATCAGT CAGGCCCTAT CTAAGAAGCT CATGGTATAC 1500
 AAGGGTCACC CAAATAGCTG AGTGCAAGTC TTGCTCATAT TTCTTTCATC TTAACCCCGC 1560
 AAACAAGAAT TAAGATGATC CCAATAAAG AAAAATTGCT CAGGAAACTG AACCTTTTTC 1620
 TGAACCAAGC ACTGTACAGA AATCTCAGGT ATTAGAGCAA CTATGGTTGA TTGAAAAGTG 1680
 TCTCAAAATC TGGGCCAAGA ATGATTGCTA GGTCCATAAG CTAATTGTGTC TGGCCTTGCC 1740
 50 ATTTACGTAA GCCAAAGAAA GTCACTCATG AGTAAACTAT AGAAAACGTT CAGACCCATC 1800
 CTGTTAGTAT GTCAAAATCA CTAAGACTGG CAGGGTATTA ACTCCATTC AGGTGACATG 1860
 GATAAAGAGC CCCATTTTTC TCACAGTGCC AGCCTCTACC TAAGGAAACC CTAGACCTTG 1920
 GAACCAAGTT CCTGGTAGGG AACTGCTGAC AGTTTCAATG CTGACAGTTG GAGCCATATC 1980
 CTCATAGTGT AAATGAAAG AAAAATAGTT GCTTTTAAAA ATGTCAGCAA GAAGGCCTGC 2040
 55 CTCATCTTAA CAAAGCAAAA AAAAATGCTT TAATTCAAAT TAAAAATCAT GATACATAAA 2100
 AAAAAAAA

SEQ ID NO:92 PDV5 Protein sequence

Protein Accession #: NP_057674

1 11 21 31 41 51
 | | | | |
 MQCQLFRITET SKAVSELNYD YICIKAGTGR PQGTPTIGLV LLVRWAIYE TELQSQPIIT

SEQ ID NO:93 PEE6 DNA SEQUENCE

Nucleic Acid Accession #: NM_002606

Coding sequence: 61-1842 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 CGCGGCGGCT GGCCTCGGGA AAGTACAGTA AAAAGTCCGA GTGCAGCCGC CGGGCGCAGG 60
 ATGGGATCCG GCTCCTCCAG CTACCGGCCC AAGGCCATCT ACCTGGACAT CGATGGACGC 120
 ATTCAGAAAG TAATCTTTCAG CAAGTACTGC AACTCCAGCG ACATCATGGA CCTGTTCTGC 180
 ATCGCCACCG GCCTGCCCTCG GAACACGACC ATCTCCCTGC TGACCACCGA CGACGCCATG 240
 70 GTCTCCATCG ACCCCACCAT GCCCGCGAAT TCAGAACGCA CTCCGTACAA AGTGAGACCT 300
 GTGGCCATCA AGCAACTCTC CGCTGGTGTG GAGGACAAGA GAACCAAG CCGTGGCCAG 360
 TCTGCTGAGA GACCACTTGT GGACAGACGG GTTGTGGGCC TGAGGACGCC CCGGAGGGAA 420
 GGAGCATTTG AAAGTGGACA GGTAGAGCCC AGGCCAGAG AGCCCCAGG CTGCTACCA 480
 GAAGGCCAGC GCATCCCTCC AGAGAGAGAA GAATTAATCC AGAGCGTGCT GCGCAGGTT 540
 GCAGAGCAGT TCTCAAGAGC ATTCAAAATC AATGAATGTA AAGCTGAAGT TGCAATCAC 600
 80 TTGCTGTGCC TAGAGAAACG CGTGGAATG GAAGGACTAA AAGTGGTGA GATTAGAAA 660

TGCAAGAGTG ACATTAAGAA GATGAGGGAG GAGCTGGCGG CCAGAAGCAG CAGGACCAAC 720
 TGCCCTGTGA AGTACAGTTT TTTGGATAAC CACAAGAAAGT TGACTCCTCG ACGCGATGTT 780
 CCCACTTACC CCAAGTACCT GCTCTCTCCA GAGACCATCG AGGCCCTGCG GAAGCCGACC 840
 TTTGACGTCT GGCTTTGGGA GCCCAATGAG ATGCTGAGCT GCCTGGAGCA CATGTACCAC 900
 5 GACCTCGGGC TGGTCAGGGA CTTACAGCATC AACCTGTGCA CCCTCAGGAG GTGGCTGTTC 960
 TGTTGCCACG ACAACTACAG AAACAACCCC TTCCACAAC TCCGGCAGCTG CTTCTGCGTG 1020
 GCCCAGATGA TGTACAGCAT GGCTCTGGCTC TGCAGTCTCC AGGAGAAAGT CTCACAAACG 1080
 GATATCTCTGA TCCTAATGAC AGCGGCCATC TGCCACGATC TGGACCATCC CGGCTACAAC 1140
 10 AACACGTACC AGATCAATGC CCGCACAGAG CTGGCGGTCC GCTACAATGA CATCTACCG 1200
 CTGGAGAACC ACCACTGCGC CGTGGCCTTC CAGATCCTCG CCGAGCCTGA GTGCAACATC 1260
 TTCTCCAACA TCCCACCTGA TGGGTTCAG CAGATCCGAC AGGGAATGAT CACATTAATC 1320
 TTGGCCACTG ACATGGCAAG ACATGCAGAA ATTATGGATT CTTTCAAAGA GAAATGGAG 1380
 AATTTTGACT ACAGCAACGA GGAGCACATG ACCCTGCTGA AGATGATTTT GATAAAATGC 1440
 15 TGTGATATCT CTAACGAGGT CCGTCCAATG GAAGTCGCAG AGCCTTGGGT GGACTGTTTA 1500
 TTAGAGGAAT ATTTTATGCA GAGCGACCGT GAGAAGTCAG AAGGCCTTCC TGTGGCACC 1560
 TTCTATGGAC GAGACAAAGT GACCAAGGCC ACAGCCGAGA TTGGGTTCAT CAAGTTTGTG 1620
 CTGATCCCAA TGTTTGAAAC AGTGACCAAG CTCTTCCCA TGGTTGAGGA GATCATGCTG 1680
 CAGCCACTTT GGGAAATCCC AGATCGCTAC GAGGAGCTGA AGCGGATAGA TGACGCCATG 1740
 20 AAAGAGTTAC AGAAGAGAC TGACAGCTTG ACGTCTGGGG CCACCGAGAA GTCCAGAGAG 1800
 AGAAGCAGAG ATGTGAAAAA CAGTGAAGGA GACTGTGCCT GAGGAAAGCG GGGGGCGTGG 1860
 CTGACGTTCT GGACGGGCTG GCCGAGCTGC GCGGGATCCT TGTGCAGGGA AGAGCTGCC 1920
 TGGGCACCTG GCACCACAAG ACCATGTTTT CTAAGAACCA TTTTGTTCAC TGATACAAA 1980
 AAAAAAAAAA A

SEQ ID NO:94 PEE6 Protein sequence

Protein Accession #: NP_002597

1 11 21 31 41 51
 30 | MGS GSSSYRP | KAIYLDIDGR | IQKVIFSKVC | NSSDIMDLFC | IATGLPRNTT | ISLLTTDDAM 60
 VSDPTMPAN | SERTPYKVRP | VAIKQLSAGV | EDKRTTSRQG | SAERPLRRR | VVGLEQPRRE 120
 GAFESQVPE | RPREPGQCVQ | EGQRIPPERE | ELIQSVLAQV | AEQFSRAFKI | NELKAEVANH 180
 LAVLEKRVEL | EGLKVVEIEK | CKSDIKMRE | ELAARSSRTN | CPCKYSFLDN | HKKLTFRPDV 240
 35 PTPYKYLSP | ETIEALRKPT | FDVWLWEPNE | MLSCLEHMYH | DLGLVRDFSI | NFVTLRRWLF 300
 CVHDNYRNP | FHNFRHCFCV | AQMMYSMVWL | CSLQEKFSQT | DILILMTAAI | CHDLDPHGYN 360
 NTYQINARTE | LAVRYNDISP | LENHHCVAFA | QILAEPECNI | FSNIPDPGFK | QIRQGMITLI 420
 LATDMARHAE | IMDSFKEKME | NFDYSNEEHM | TLLKMILIKC | CDISNEVRPM | EVAEPWVDCL 480
 LEEYFMQSDR | EKSEGLEPVA | FMDRDKVTKA | TAQIGFIKVF | LIPMFETVTK | LFPMVEETML 540
 40 QPLWESDRY | EELKRIDDAM | KELQKKTDSL | TSGATEKSRE | RSRDVKNSEG | DCA

SEQ ID NO:95 PEG4 DNA SEQUENCE

Nucleic Acid Accession #: none

Coding sequence: 41-559 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 45 | CAGTCACAGG | CGAGAGCCYT | GGGATGCACC | GGCCAGAGGC | ATGCTGCTGC | TGCTCACGCT 60
 TGCCCTCCTG | GGGGGCCCCA | CCTGGGCAGG | GAAGATGTAT | GGCCCTGGAG | GAGGCAAGTA 120
 50 TTTAGCAGCC | ACTGAAAGCT | ACGACCATGA | AATCAGAGG | CTGCGGGTGT | CTGTAGGTCT 180
 TCCTCTGGTG | AAAAGTGTCC | AGGTGAAACT | TGGAGACTCC | TGGGACGTGA | AACTGGGAGC 240
 CTTAGGTGGG | AATACCCAGG | AAGTCAACCT | GCAGCCAGGC | GAATACATCA | CAAAAGTCTT 300
 TGTCCGCTTC | CAAGCTTTC | TCCGGGGTAT | GGTCATGTAC | ACCAGCAAGG | ACCGCTATTT 360
 CTATTTTGGG | AAGCTTGATG | GCCAGATCTC | CTCTGCCTAC | CCCAGCCAAG | AGGGGCAGGT 420
 55 GCTGGTGGGC | ATCTATGGCC | AGTATCAACT | CCTTGGCATC | AAGAGCATTG | GCTTTGAATG 480
 GAATTATCCA | CTAGAGGAGC | CGACCACTGA | GCCACCAGTT | AATCTCACAT | ACTCAGCAAA 540
 CTCACCCGTG | GGTGCGTAGG | GTGGGGTATG | GGGCCATCCG | AGCTGAGGCC | ATCTGTGTGG 600
 TGGTGGCTGA | TGGTACTGGA | GTAACGTAGT | CGGGACGCTG | AATCTGAATC | CACCAATAAA 660
 TAAAGCTTCT | GCAGAATCAG | TGAATAAAAA A

SEQ ID NO:96 PEG4 Protein sequence

Protein Accession #: FGENESH predicted

1 11 21 31 41 51
 65 | MLLLLTLALL | GGPTWAGKMY | GPGGGKYFST | TEDYDHEITG | LRVSVGLLLV | KSVQVKLGDS 60
 WDKLGALGG | NTQEVTLQPG | EYITKVVFVA | QAFLRGMVMY | TSKDRYFYFG | KLDGQISSAY 120
 PSQEGQVLVG | IYQYQLLGI | KSIQFEWNYF | LEEPTTEPFV | NLTYSANSEFV | GR

SEQ ID NO:97 PEL9 DNA SEQUENCE

Nucleic Acid Accession #: NM_006953

Coding sequence: 33-896 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 75 | CCGTTCCGCG | CTCTGGCGGC | TCCTCCCGGG | CGATGCCTCC | GCTCTGGGCC | CTGCTGGCCC 60
 TCGGCTGCCT | GCGGTTCCGG | TCGGCTGTGA | ACCTGCAGCC | CCAACTGGCC | AGTGTGACTT 120
 TCGCCACCAA | CAACCCACAC | CTTACCACTG | TGGCCTTGGA | AAAGCCTCTC | TGCATGTTTG 180
 80 ACAGCAAAGA | GGCCTTCACT | GGCACCCACG | AGGTCTACCT | GTATGTCTCG | GTCGACTCAG 240
 CCATTTCAG | GAATGCCTCA | GTGCAAGACA | GCACCAACAC | CCCACTGGGC | TCAACGTTC 300

TACAAACAGA GGGTGGGAGG ACAGGTCCCT ACAAAGCTGT GGCCCTTGAC CTGATCCCCT 360
 GCAGTGACCT GCCCAGCCTG GATGCCATTG GGGATGTGTC CAAGGCCTCA CAGATCCTGA 420
 ATGCCCTACCT GGTACAGGCTG GGTGCCAACG GGACCTGCCT GTGGGATCCC AACTTCCAGG 480
 GCCTCTGTAA CGCACCCCTG TCGGCAGCCA CGGAGTACAG GTTCAAGTAT GTCCTGGTCA 540
 ATATGTCCAC GGGCTTGGTA GAGGACCAGA CCCTGTGGTC GGACCCATC CGCACCAACC 600
 AGCTCACCCT ATACTCGACG ATCGACACGT GGCCAGGCCG GCGGAGCGGA GGCATGATCG 660
 TCATCACTTC CATCTTGGGC TCCCTGCCCT TCTTCTACT TGTGGGTTT GCTGGCGCCA 720
 TTGCCCTCAG CCTCGTGGAC ATGGGGAGTT CTGATGGGGA AACGACTCAC GACTCCCAAA 780
 TCACTCAGGA GGCTGTTCCT AAGTCGCTGG GGGCCTCGGA GTCTTCTTAC ACGTCCGTGA 840
 ACCGGGGGCC GCCACTGGAC AGGGCTGAGG TGTATTCAG CAAGCTCCAA GACTGAGCCC 900
 AGCACCAACC CTGGGACAGA GCATCTCTCT CTCTGGCCTT GCCCAGGCC CTGCAGCGGT 960
 GGTGTGCACA CCCTGACTTC AGGGAAGGTG AAACAGGGCT TGTCCCTCCA ACTGCAGGAA 1020
 AACCCTTAAT AAAATCTTCT GATGAGTTCT AAAAAAAA

SEQ ID NO:98 PEL9 Protein sequence
 Protein Accession #: NP_008884

1 11 21 31 41 51
 | | | | |
 MPPLWALLAL GCLRFGSAVN LQPQLASVTF ATNNPFLTTV ALEKPLCMFD SKEALTGTHE 60
 VYLYVLVDSA ISRNASVQDS TNTPLGSTFL QTEGGRTGPY KAVAFDLIPC SDLPSLDAIG 120
 DVSKASQILN AYLVVRVGANG TCLWDNPFQ LCNAPLSAAT EYRFKYVLVN MSTGLVEDQT 180
 LWSDFPRTNQ LTPYSTIDTW PRRSGGMIV TTSILGSLPF FLLVGFAGAI ALSLVDMSGSS 240
 DGETTHDSQI TQEAIVPKSLG ASESSYTSVN RGPPLDRAEV YSSKLQD

SEQ ID NO:99 PEN1 DNA SEQUENCE

Nucleic Acid Accession #: NM_012391
 Coding sequence: 416-1423 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 GTCTGACTTC CTCCAGCAGC ATTCTGTCAC TCTGCCGTGT CCACACTGCC CCACAGACCC 60
 AGTCCCTCCAA GCCTGTGTGC AGCTCCCTGC AAGCCCTCA GGTGGGCCT TGCCACGGTG 120
 CCAGCAGGCA GCCCTGGGCT GGGGGTAGGG GACTCCCTAC AGGCACGCAG CCCTGAGACC 180
 TCAGAGGGCC ACCCTTTGAG GGTGGCCAGG CCCCAGTGG CCAACCTGAG TGCTGCCCTCT 240
 GCACACAGCC CTGCTGGGCC CTGGTTCGCG TGGCCCCCA GATGCCCTGGC TGAGACACGC 300
 CAGTGGCCTC AGTGTCCAC ACCTCTTCCC GGCCCTGAA GTTGGCCTG CAGCAGACAG 360
 CTCCCTGGGC ACCAGGCAGC TAACAGACAC AGCCGCCAGC CCAACAGCA GCGGCATGGG 420
 CAGCGCCAGC CCGGGTCTGA GCAGCGTATC CCCCAGCCAC CTCTGCTGTC CCCCAGACAC 480
 GGTGTGCGCG ACAGGCTTGG AGAAGGCGGC AGCGGGGGCA GTGGGTCTCG AGAGACGGGA 540
 CTGGAGTCCC AGTCCACCCG CCACGCCCGA GCAGGGCCTG TCCGCCTTCT ACCTCTCCTA 600
 CTTTGACATG CTGTACCCCTG AGGACAGCAG CTGGGCAGCC AAGGCCCCTG GGGCCAGCAG 660
 TGGGAGGAG CCACCTGAGG AGCCTGAGCA GTGCCCGGTC ATTGACAGCC AAGCCCCAGC 720
 GGGCAGCCTG GACTTGGTGC CCGGCGGGCT GACCTTGGAG GAGCACTCGC TGGAGCAGGT 780
 GCAGTCCATG GTGGTGGGCG AAGTGCTCAA GGACATCGAG ACGGCCTGCA AGCTGCTCAA 840
 CATCACCCGA GATCCCATGG ACTGGAGCCC CAGCAATGTG CAGAAGTGGC TCCTGTGGAC 900
 AGAGCACCAA TACCGGCTGC CCCCATGGG CAAGGCCTTC CAGGAGCTGG CGGGCAAGGA 960
 GCTGTGCGCC ATGTGCGAGG AGCAGTTCCT CCAGCGCTCG CCCCTGGGTG GGGATGTGCT 1020
 GCACGCCAC CTGGACATCT GGAAGTCAGC GGCCCTGGATG AAAGAGCGGA CTTCACTGG 1080
 GGCATTCAC TACTGTGCTT CGACAGTGA GGAGAGCTGG ACCGACAGCG AGGTGGACTC 1140
 ATCATGCTCC GGGCAGCCCA TCCACCTGTG GCAGTTCCTC AAGGAGTTGC TACTCAAGCC 1200
 CCACAGCTAT GGCCGCTTCA TTAGTGGCT CAACAAGGAG AAGGCATCT TCAAAATTGA 1260
 GGACTCAGCC CAGGTGGGCC GGCTGTGGGG CATCCGCAAG AACCTCCCG CCATGAACCTA 1320
 CGACAAGCTG AGCCGCTCCA TCCGCCAGTA TTACAAGAAG GGCATCATCC GGAAGCCAGA 1380
 CATCTCCAG CGCCTCGTCT ACCAGTTCGT GCACCCCATC TGAGTGCCTG GCCCAGGGCC 1440
 TGAACCCCGC CTTACAGGGC CTCTCTCCTG CCTGCCCTGC CTCAGCCAGG CCCTGAGATG 1500
 GGGGAAAACG GGCAGTCTGC TCTGCTGCTC TGACCTTCCA GAGCCCAAGG TCAGGGAGGG 1560
 GCAACCAACT GGGCAGGGG GATATGGGTC CTCTGGGGCC TTCGGGACCA TGGGGCAGGG 1620
 GTGCTTCCCT CTCAGGCCCA GCTGCTCCCC TGGAGGACAG AGGGAGACAG GGCTGCTCCC 1680
 CAACACCTGC CTCTGACCCC AGCATTTCCA GAGCAGAGCC TACAGAAGGG CAGTGACTCG 1740
 ACAAAAGCCA CAGGCAGTCC AGGCCTCTCT CTGCTCCATC CCCCTGCCCT CCATTCCTGA 1800
 CCACACCTGG CATGGTGCAG GGAGACATCT GCACCCCTGA GTTGGGCAGC CAGGAGTGCC 1860
 CCCGGGAATG GATAATAAAG ATACTAGAGA ACTG

SEQ ID NO:100 PEN1 Protein sequence
 Protein Accession #: NP_036523

1 11 21 31 41 51
 | | | | |
 MGSASPLGSS VSPSHLLLP DTVSRTGLEK AAAGAVGLER RDWSFSPFAT PEQGLSAFYL 60
 SYFDMLYPED SSWAAKAFGA SSREEPPEEP EQCFVIDSQA PAGSLDLVPG GLTLEHSLE 120
 VQVSMVVEV LKDIETACKL LNITADPMDW SPSNVQKWLW WTEHQYRLFP MGKAFQELAG 180
 KELCAMSEEQ FRQRSPLGGD VLHAHLDIWK SAAMWKERTS PGAIHYCAST SEESWTDSEV 240
 DSSCSGQPIH LWQFLKELL LKPHSYGRFIR WLNKEKGIFK IEDSAQVARL WGIRKNRPAM 300
 NYDKLSRSIR QYKKGIIRK PDISQRLVYQ FVHPI

SEQ ID NO:101 PEN3 DNA SEQUENCE

Nucleic Acid Accession #: NM_000742
 Coding sequence: 555-2144 (underlined sequences correspond to start and stop codons)

	1	11	21	31	41	51	
	GAGAGAACAG	CGTGAGCCTG	TGTGCTTGTG	TGCTGAGCCC	TCATCCCCCTC	CTGGGGCCAG	60
5	GCTTGGGTTT	CACCTGCAGA	ATCGCTTGTG	CTGGGCTGCC	TGGGCTGTCC	TCAGTGGCAC	120
	CTGCATGAAG	CCGTTC TGCG	TGCCAGAGCT	GGACAGCCCC	AGGAAAACCC	ACCTCTCTGC	180
	AGAGCTTGCC	CAGCTGTCCC	CGGGAAGCCA	AATGCCTCTC	ATGTAAGTCT	TCTGCTCGAC	240
	GGGGTGTCTC	CTAAACCCTC	ACTCTTCAGC	CTCTGTTTGA	CCATGAAATG	AAGTGACTGA	300
	GCTCTATTCT	GTACCTGCCA	CTCTATTCTC	GGGGTGACTT	TTGTCAAGTG	CCCAGAATCT	360
10	CCAAGCCAGG	CTGGTTCTCT	GCATCCTTTC	AATGACCTGT	TTTCTTCTGT	AACCACAGGT	420
	TCGGTGGTGA	GAGGAAGCCT	CGCAGAATCC	AGCAGAATCC	TCACAGAATC	CAGCAGCAGC	480
	TCTGCTGGGG	ACATGGTCCA	TGGTGCAACC	CACAGCAAAG	CCCTGACCTG	ACCTCCTGAT	540
	GCTCAGGAGA	AGCCATGGGC	CCCTCCTGTC	CTGTGTTTCT	GTCCCTTCACA	AAGCTCAGCC	600
	TGTGGTGGCT	CCTTCTGACC	CCAGCAGGTG	GAGAGGAAAG	TAAGCGCCCA	CCTCCAGGG	660
15	CTCCTGGAGA	CCCACCTCTC	TCTCCAGTTC	CCACGGCATT	GCCCGCAGGA	GGCTCGCATA	720
	CCGAGACTGA	GACCCGGCTC	TTCAAACACC	TCTTCCGGGG	CTACAACCCG	TGGGCGCGCC	780
	CGGTGCCCAA	CACCTTCAGC	GTGGTGATTC	TGCGCTTTGG	ACTGTCCATC	GCTCAGCTCA	840
	TCGATGTGGA	TGAGAAGAAC	CAATGATGA	CCACCAACGT	CTGGCTAAAA	CAGGAGTGA	900
	GCGACTACAA	ACTGCGCTGG	AACCCCGCTG	ATTTTGGCAA	CATCACATCT	CTCAGGGTCC	960
20	CTTCTGAGAT	GATCTGGATG	CCCGACATTC	TTCTCTACAA	CAATGCAGAT	GGGGAGTTTG	1020
	CAGTGACCCA	GATGACCAAG	GCCACCTCTC	TCTCCACGGG	CAGTGTGCAC	TGGGTGCCCC	1080
	CGGCCATCTA	CAAGAGCTCC	TGCAGCATCG	ACGTCACTTT	CTTCCCTTTC	GACCAAGCAGA	1140
	ACTGCAAGAT	GAAGTTTGGC	TCCTGGACTT	ATGACAAGGC	CAAGATCGAC	CTGGAGCAGA	1200
	TGGAGCAGAC	TGTGGACCTG	AAGGACTACT	GGGAGAGCGG	CGAGTGGGCC	ATCGTCAATG	1260
25	CCACGGGCAC	CTACAACAGC	AAGAAGTACG	ACTGCTGCGC	CGAGATCTAC	CCCGACGCTCA	1320
	CCTACGCCTT	CGTCACTCCG	CGGCTGCCCG	TCTTCTACAC	CATCAACCTC	ATCATCCCTC	1380
	GCCTGCTCAT	CTCCTGCCTC	ACTGTGCTGG	TCTTCTACCT	GCCCTCCGAC	TGCGGCGAGA	1440
	AGATCAGCTG	GTGCATTTTC	GTGCTGCTGT	CACCTACCGT	CTTCTCTGCT	CTCATCACTG	1500
	AGATCATCCG	GTCCACCTCG	CTGGTCATCC	CGCTCATCGG	CGAGTACCTG	CTGTTCACCA	1560
30	TGATCTTCGT	CACCCCTGTC	ATCGTCATCA	CCGTCTTCGT	GCTCAATGTG	CACCAACGCT	1620
	CCCCCAGCAC	CCACACCATG	CCCCACTGGG	TGCGGGGGGC	CCTTCTGGGC	TGTGTGCCCC	1680
	GGTGGCTTCT	GATGAACCGG	CCCCCACCAC	CGGTGGAGCT	CTGCCACCCC	CTACGCTTGA	1740
	AGCTCAGCCC	CTCTTATCAC	TGGCTGGAGA	GCAACGTGGA	TGCCGAGGAG	AGGGAGGTGG	1800
	TGGTGGAGGA	GGAGGACATG	TGGGCATGTG	CAGGTCTATG	GGCCCCCTCT	GTGGGCACCC	1860
35	TCTGCAGCCA	CGGCCACCTG	CACCTTGGGG	CCTCAGGTCC	CAAGGCTGAG	GCTCTGTGTC	1920
	AGGAGGGTGA	GCTGCTGCTA	TCACCCACCA	TGCAGAAGGC	ACTGGAAGGT	GTGCACTACA	1980
	TTGCCGACCA	CCTGCCGTCT	GAGGATGCTG	ACTCTTCTGT	GAAGGAGGAC	TGGAAGTATG	2040
	TTGCCATGGT	CATCGACAGG	ATCTTCTCTC	GGCTGTTTAT	CATCGTCTGC	TTCTTGGGGA	2100
40	CCATCGGCTT	CTTCTCGCCT	CGGTTCCTAG	CTGGAATGAT	CTGACTGCAC	CTCCCTCGAG	2160
	CTGGCTCCCA	GGGCAAAGGG	GAGGGTTCTT	GGATGTGGAA	GGGCTTTGAA	CAATGTTTAG	2220
	ATTTGGAGAT	GAGCCCAAAG	TGCCAGGGAG	AACAGCCAGG	TGAGGTGGGA	GGTTGGAGAG	2280
	CCAGGTGAGG	TCTCTCTAAG	TCAGGCTGGG	GTGTAAGTTT	GGAGTCTGTC	CGAGTTTGCA	2340
	GGTGCTGTAG	CTGTATGGTC	CAGCAGGGGA	GTAATAAGGG	CTCTTCCGGA	AGGGGAGGAA	2400
	GCGGGAGGCA	GGCCTGCACC	TGATGTGGAG	GTACAGGCAG	ATCTTCCCTA	CCGGGGAGGG	2460
45	ATGGATGGTT	GGATACAGGT	GGCTGGGCTA	TTCCATCCAT	CTGGAAGCAC	ATTTGAGCGT	2520
	CCAGGCTTCT	CCTTGACGTC	ATTCCTCTCC	TTCTTGTCTG	CAAAATGGCT	CTGCACCAGC	2580
	CGGCCCCAG	GAGGCTCTGC	AGAGCTGAGA	GCCATGGCCT	GCAGGGGGTC	CATATGTCCC	2640
	TACGCGTGCA	GCAGGCAAAC	AAGA				

SEQ ID NO:102 PEN3 Protein sequence

Protein Accession #: NP_000733

	1	11	21	31	41	51	
55	MGSPCVFLS	FTKLSLWLL	LTPAGGEEAK	RPPPRAPGDP	LSSPSPTALP	QGGSHETED	60
	RLFKHLFRGY	NRWARVPVNT	SDVVIIVRFL	SIAQLIDVDE	KNQMMTTNVW	LKQWSDYKL	120
	RWNPADEFNI	TSLRVPSEMI	WIPDIVLYNN	ADGEFAVTHM	TKAHLFSTGT	VHWVPPAIYK	180
	SSCSIDVTFF	PFDDQNCKMK	FGSWTYDKAK	IDLEQMEQTV	DLKDYWESGE	WAIIVNATGY	240
	NSKKYDCCAE	IYFDVITYAFV	IRRLPLFYTI	NLIIPCLLIS	CLTVLVFVLP	SDCGEKITLC	300
60	ISVLLSLTVF	LLLITELIPS	TSLVIPLIGE	YLLFTMIFVT	LSIVITVFVL	NVHHRSPSTH	360
	TMPHWVRGAL	LGCVPWRLLM	NRPPPPVELC	HPLRLKLSPS	YHWLESNVDA	EEREVVVEEE	420
	DRWACAGHVA	PSVGTLCSHG	HLHSGASGPK	AEALLQEGEL	LLSPHMOKAL	EGVHYIADHL	480
	RSEDADSSVK	EDWKYVAMVI	DRIFLWLFII	VCFLGTITGLF	LPPFLAGMI		

SEQ ID NO:103 PEU4 DNA SEQUENCE

Nucleic Acid Accession #: NM_018670

Coding sequence: 87-893 (underlined sequences correspond to start and stop codons)

	1	11	21	31	41	51	
70	CACGAGGCTG	GAAGGGGCCA	CTTCACACCT	CGGGCTCGGC	ATAAAGCGGC	CGCCGGCCGC	60
	CGGCCCCAGC	ACGCGCCGCC	GCTGCCATGG	CCCAGCCCTT	GTGCCCCCGC	CTCTCCGAGT	120
	CCTGGATGCT	CTCTCGCGCC	TGGGGCCCAA	CTCGGCGGCC	GCCGCGCTCC	GACAAGGACT	180
	GCGGCGGCTC	CCTCGTCTCG	TCCCCAGACT	CATGGGGCAG	CACCCCAGCC	GACAGCCCGC	240
75	TGGCGAGCCC	CGCGCGGCCA	GGCACCCCTC	GGGACCCCGC	CGCCCCCTCC	GTAGGTAGGC	300
	GCGGCGCGCG	CAGCAGCCGC	CTGGGCAGCG	GGCAGAGGCA	GAGCGCCAGT	GAGCGGGAGA	360
	AACTGCGCAT	GCGCACGCTG	GCCCGCGCCC	TGCACGAGCT	GCGCGGCTTT	CTACCGCCGT	420
	CCGTGGCGCC	CGCGGGCCAG	AGCCTGACCA	AGATCGAGAC	GCTGCGCCTG	GCTATCCGCT	480
	ATATCGGCCA	CCTGTGCGCC	GTGCTAGGCC	TCAGCGAGGA	GAGTCTCCAG	CGCCGGTGCC	540
80	GGCAGCGCGG	TGACGCGGGG	TCCCTTCGGG	GCTGCCCGCT	GTGCCCCGAC	GACTGCCCGG	600
	CGCAGATGCA	GACACGGACG	CAGGCTGAGG	GGCAGGGGCA	GGGGCGCGGG	CTGGGCCTGG	660

TATCCGCCGT CCGCGCCGGG GCGTCCGTTG GATCCCCGCC TGCCTGCCCC GGAGCCCGAG 720
 CTGCACCCGA GCCCGCGGAC CCGCCTGCGC TGTTCCGCGA GCGCGCGTGC CCGGAAGGGC 780
 AGGCGATGGA GCCAAGCCCA CCGTCCCGCG TCCTTCCGGG CGACGTGCTG GCTCTGTTGG 840
 AGACCTGGAT GCCCTCTCG CCTCTGGAGT GGCTGCCCTGA GGAGCCCAAG TGACAAGGGA 900
 CAACTGACGC CGTCTCTGTG AGCACCAGAG CTTTTTGGCC TCAGCACCTT CGAAGTGGTT 960
 CCTTGGCAGA CTGCTTTTCC TGGAAAGAGG CACGGGCGAT CCCGACGGGG GCATTCTCTG 1020
 GGGTGAGAGC CTGCCCCACC GCGGCGGCCC TTCTCAGCCC CTCCCCTCCAT GGAGGGACCC 1080
 ATAGGGCTAG ACACCTTTGAG GCAAGCAGGA GGCCTCTGCCT AATGTGAATT TATTTATTTC 1140
 TGAATAAACT GTACTGTGT CAAAAA AAAA A A

SEQ ID NO:104 PEU4 Protein sequence

Protein Accession #: NP_061140

1 11 21 31 41 51
 MAQPLCPPLS ESWMLSAAWG PTRRPPPSDK DCGRSLVSSP DSWGSTPADS FVASPARPGT 60
 LRDPRAPSVG RRGARSSRLG SGQRQSASER EKLRMRTLAR ALHELRRFLP FSVAPAGQSL 120
 TKIETLRLAI RYIHLSSAVL GLSEESLQRR CRQRGDAGSP RGCPLCPDDC PAQMQRTRTQA 180
 EGQGGQGRGLG LVSAVRAGAS WGSPPACPGA RAAPEPRDPP ALFAEAACPE GQAMEPSPPS 240
 PLLPGDVLAL LETWMLPSPL EWLPEEPK

SEQ ID NO:105 PEU5 DNA SEQUENCE

Nucleic Acid Accession #: NM_017636

Coding sequence: 324-3374 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 CCACGGAGAA GCCCACCGAT GCCTACGGAG AGCTGGACTT CACGGGGGCC GGCCGCAAGC 60
 ACAGCAATTT CCTCCGGCTC TCTGACCGAA CGGATCCAGC TGCAGTTTAT AGTCTGGTCA 120
 CACGCACATG GGGCTTCCGT GCCCGAACC TGGTGGTGTG AGTGTGGGG GGATCGGGGG 180
 GCGCCGTCCT CCAGACCTGG CTGCAGGACC TGCTGCGTCG TGGGCTGGTG CGGGCTGCC 240
 AGAGCACAGG AGCCTGGATT GTCACTGGGG GTCTGCACAC GGGCATCGCG CGGCATGTTG 300
 GTGTGGCTGT ACGCGACCAT CAGATGGCCA GCACTGGGG CACCAAGGTG GTGGCATGG 360
 GTGTGGCCCC CTGGGGTGTG GTCCGGAATA GAGACACCTT CATCAACCCC AAGGGCTCGT 420
 TCCCTGCGAG GTACCGGTGG CGCGGTGACC CGGAGGACGG GGTCCAGTTT CCCCTGGACT 480
 ACAACTACTC GGCCTTCTTC CTGGTGGACG ACGGCACACA CGGCTGCC TG GGGGGCGAGA 540
 ACCGCTTCCG CTTGCGCCTG GAGTCTCTACA TCTCACAGCA GAAGACGGGC GTGGGAGGGA 600
 CTGGAATTGA CATCCCTGTG CTGCTCTCTC TGATTGATGG TGATGAGAAG ATGTTGACGC 660
 GAATAGAGAA CGCCACCCAG GCTCAGCTCC CATGTCTCCT CGTGGCTGGC TCAGGGGGAG 720
 CTGCGGACTG CCTGGCGGAG ACCCTGGAAG ACACCTCTGGC CCCAGGGAGT GGGGGAGCCA 780
 GGCAAGGCGA AGCCCGAGAT CGAATCAGGC GTTTCTTTCC CAAAGGGGAC CTTGAGGTCC 840
 TGCAGGCCCA GGTGAGAGAG ATTATGACCC GGAAGGAGCT CCTGACAGTC TATCTCTCTG 900
 AGGATGGGTC TGAGGAATTC GAGACATAG TTTTGAAGGC CTTGTGAAG GCCTGTGGGA 960
 GCTCGGAGGC CTCAGCTTAC CTGGATGAGC TGCGTTTGGC TGTGGCTTGG AACCGCGTGG 1020
 ACATTCGCCA GAGTGAATCT TTTCCGGGGG ACATCCAATG CGGCTCTCTC CATCTCGAAG 1080
 CTTCCCTCAT GGACGCTCTG CTGAATGACC GGCCTGAGTT CGTGCCTTGG CTCAATTTCC 1140
 ACGGCTCAG CCTGGGCCAC TTCTTGACCC CGATGCGCCT GGCCCAACTC TACAGCGCGG 1200
 CGCCCTCCAA CTCGCTCATC CGCAACCTTT TGGACCAGGC GTCCACAGC GCAGGCACCA 1260
 AAGCCCAAGC CTTAAAGGG GAGCTGCGG AGCTCCGGCC CCTTGACGTG GGGCATGTGC 1320
 TGAGGATGCT GCTGGGGAAG ATGTGCGCGC CGAGGTACCC CTCCGGGGGC GCCTGGGACC 1380
 CTCACCCAGG CCAGGGCTTC GGGGAGAGCA TGTATCTGCT CTCGGACAAG GCCACCTCGC 1440
 CGCTCTCGCT GATGCTGGC CTCGGGCAGG CCCCTTGGAG CGACCTGCTT CTTTGGGCAC 1500
 TGTGTCTGAA CAGGGCACAG ATGGCCATGT ACTTCTGGGA GATGGGTTC AATGCAGTTT 1560
 CCTCAGCTCT TGGGCTCTGT TTGCTGCTCC GGGTGTGGC ACGCTGGAG CTTGACGCTG 1620
 AGGAGGCAGC ACGGAGGAAA GACCTGGCGT TCAAGTTTGA GGGGATGGGC GTTGACCTCT 1680
 TTGGCGAGTG CTATCGCAGC AGTGAGGTGA GGGCTGCCCG CCTCCTCCTC CGTCCGCTGCC 1740
 CGCTCTGGGG GGATGCCACT TGCCCTCAGC TGGCCATGCA AGCTGACGCC CGTGCCTTCT 1800
 TTGCCACGGA TGGGGTACAG TCTCTGCTGA CACAGAAGTG GTGGGGAGAT ATGGCCAGCA 1860
 CTACACCCAT CTGGGCCCTG GTTCTCGCCT TCTTTTGGCC TCCACTCATC TACACCCGCC 1920
 TCATACCTTT CAGGAAATCA GAAGAGGAGC CCACACGGGA GGAGCTAGAG TTTGACATGG 1980
 ATAGTGTAT TAATGGGGAA GGGCCTGTGC GGACGGCGGA CCCAGCCGAG AAGACGCCCG 2040
 TGGGGGTCCC CCGCCAGTCG GGCCTGCGG GTTGTCTGCG GGGCCGCTGC GGGGGGCGCC 2100
 GGTGCTACG CCGCTGGTTC CACTTCTGGG GCGCGCCGGT GACCATCTTC ATGGGCAACG 2160
 TGGTCAGCTA CTTGCTGTTT TTGCTGCTTT TCTCGCGGGT GCTGCTCGTG GATTTCAGC 2220
 CGGCGCCGCC CGGCTCCCTG GAGCTGCTGC TCTATTCTG GGCCTTCACG CTGCTGTGCG 2280
 AGGAAGTGG CCGAGGCCCT AGCGGAGGCG GGGGACGCT CGCCAGCGGG GGGCCCGGGC 2340
 CTGGCCATGC CTCATGAGC CAGCGCTGCG GCTTCTACCT CGCCGACAGC TGGAAACAGT 2400
 GCGACCTAGT GGCCTCACC TGCTTCTTCC TGGGCGTGGG CTGCGCGCTG ACCCCGGGTT 2460
 TGTACCACCT GGGCCGCACT GTCTCTGCA TCGACTTCAT GGTTTTCACG GTGCGGCTGC 2520
 TTTACATCTT CACGGTCAAC AACAAGCTGG GGGCCAGAT CGTCACTGTT AGCAAGATGA 2580
 TGAAGGACGT GTTCTTCTTC CTCTTCTTCC TCGGCGTGTG GCTGGTAGCC TATGGCGTGG 2640
 CCACGGAGGG GCTCCTGAGG CCACGGGACA GTGACTTCCC AAGTATCTCT GCGCCGCTCT 2700
 TCTACCGTCC CTACCTGACG ATCTTCGGGC AGATTCCCA GAGGACATG GACGTGGCCC 2760
 TCATGGAGCA CAGCAACTGC TCGTCCGAGC CCGGCTTCTG GGCACACCTT CCTGGGGCCC 2820
 AGGCGGGCAC CTGCTCTTCC CAGTATGCCA ACTGGCTGGT GGTGCTGCTC CTCGTCTATC 2880
 TCGTCTGCTG GGCACACATC CTGCTGGTCA ACTTGTCTAT TGCCATGTTT AGTTACACAT 2940
 TCGGCAAGT ACAGGGCAAC AGCGATCTCT ACTGGAAGGC GCAGCGTTAC CGCCTCATCC 3000
 GGGAAATTCCA CTCTCGGCCC GCGCTGGCCC CGCCCTTTAT CGTCATCTCC CACTTGGGCC 3060
 TCCTGCTCAG GCAATTGTGC AGGCGACCCC GGAGCCCCCA CGCGCTCTCC CCGGCCCTCG 3120
 AGCATTTCCG GGTATTACCTT TCTAAGGAAG CCGAGCGGAA GCTGCTAACG TGGGAATCGG 3180

TGCATAAGGA GAACCTTTCG CTGGCACGCG CTAGGGACAA GCGGGAGAGC GACTCCGAGC 3240
 GTCTGGAGCG CACGTCGCCAG AAGGTGGACT TGGCACTGAA ACAGCTGGGA CACATCCGCG 3300
 AGTACGAACA GCGCCTGAAA GTGCTGGAGC GGGAGGTCCA GCAGTGTAGC CCGCTCCTGG 3360
 GGTGGGTGAC GTAGGCCGTT AGCAGCTCTG CCATGTTGCC CTCAGGTGGG CCGCCACCCC 3420
 TTGACCTGCA TGGGTCCAAA GAGTGAGCCA TGCTGGCGGA TTTTAAGGAG AAGCCCCCAC 3480
 AGGGGATTTT GCTCTTAGAG TAAGGCTCAT GTGGGCTCG GCCCCGCAC CTGGTGGCCT 3540
 TGTCTTTGAG GTGAGCCCCA TGTCCATCTG GGCCACTGTC AGGACCACCT TTGGGAGTGT 3600
 CATCCTTACA AACCACAGCA TGCCCGGCTC CTCCAGAAC CAGTCCAGC CTGGGAGGAT 3660
 CAAGGCTGG ATCCCGGGCC GTTATCCATC TGGAGGCTGC AGGGTCTCTG GGGTAACAGG 3720
 GACCACAGAC CCCTCACCAC TCACAGATTC CTCACACTGG GGAATAAAG CCATTTTACA 3780
 GGAAAAAA AAAAAAAA AAAAAAAA

SEQ ID NO:106 PEU5 Protein sequence

Protein Accession #: NP_060106

1 11 21 31 41 51
 MASTGGTKVV AMGVAPWGVV RNRDTLINPK GSFPARYRWR GDPEDGVQFP LDYNSAFFFL 60
 VDDGTHGCLG GENRFLRLLE SYISQKQTVV GGTGIDIPVL LLLIDGDEKM LTRINATQA 120
 QLPCLLVAGS GGAADCLABT LEDTLAPGSG GARQGEARDR IRRFFPKGDL EVLQAQVERI 180
 MTRKELLTVY SSEDGSEEFV TIVLKALVKA CGSSEASAYL DELRLAVAWN RVDIAQSELF 240
 RGDIQWRSFH LEASLMDALL NDRPEFVRLI ISHGLSLGHF LTPMRLAQLY SAAPSNSLIR 300
 NLLDQASHSA GTKAPALKGG AAELRPPDVG HVLRLMLGKM CAPRYPSSGA WDPHFGQSGF 360
 ESMYLLSDKA TSPLSLDAGL GQAPWSDLLL WALLLNRAQM AMYFWEMGSN AVSSALGACL 420
 LLRVMARLEP DAEEAARRKD LAFKPEGMGV DLFGEYRVS EVRAARLLLR RCPWLGDATE 480
 LQLAMQADAR AFFAQDGVQS LLTKQWGWDM ASTTPIWALV LAFFCPLIY TRLITFRKSE 540
 BEPTRELELF DMSVINGBG PVGTADPAEK TPLGVPRQSG RPGCCGGRCG GRRCLRRWFH 600
 FWGAPVTIFM GNVVSYLLFL LLFSRVLLVD FQPAPPGSLE LLLYFWAFTL LCEELRQGLS 660
 GGGGSLASGG PGPGHASLSQ RLRLYLADSW NQCDLVALTC FLLGVGCRIT PGLYHLGRTV 720
 LCIDFMVFTV RLLHFTVWVK QLGPKIVIVS KMMKDVFFFL FFLGVWLVAV GVATEGLLRP 780
 RDSDFPSILR RVFYRPLYLI FGQIPQEDMD VALMEHSNCS SEPGFWAHP GAQAGTCVSQ 840
 YANWLVLVLL VIFLLVANIL LVNLLIAMFS YTFGKVQONS DLYWKAQRVR LIREFHSRPA 900
 LAPPFIVISH LRLLLRQCLR RPRSPQSPS ALEHFRVYLS KEAERKLLTW ESVHKENFLL 960
 ARARBKRESL SERLERTSQK VDLALKQLGH IREYEQRLKV LEREVQQCSR VLGWVT

SEQ ID NO:107 PEW3 DNA SEQUENCE

Nucleic Acid Accession #: NM_005982

Coding sequence: 276-1130 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 GGTAGCAGCA TCCACCGGGC GGGAGGTCGG AGGCAGCAAG GCCTTAAAGG CTACTGAGTG 60
 CGCCGGCGGT TCCGTGTCCA GAACCTCCCC TACTCCTCCG CCTTCTCTTC CTGGCCGCC 120
 CACCGCCAAG TTCCGACTCC GGTTTTCGCC TTTGCAAAGC CTAAGGAGGA GGTTAGGAAC 180
 AGCCGCGCCC CCTTCCCTGC GGGCGCGGCC CCTGCTCTCT CGGCTCTGCT CCCTGCCGCG 240
 TGCGCCTGGG CCGTGCAGCC CGGCAGGCGC CAGCCATGTC GATGCTGCCG TCGTTTGGCT 300
 TTACGCAGGA GCAAGTGGCG TGCGTGTGCG AGGTTCGTCG GCAAGGCGGA AACCTGGAGC 360
 GCCTGGGCAG GTTCTGTGGG TCACCTGCCG CCTGCGACCA CTGCAACAAG AACGAGAGCG 420
 TACTCAAGGC CAAGCGGGTG GTCCGCTTCC ACCGCGGCAA CTTCCGTGAG CTCTACAAGA 480
 TCCTGGAGAG CCACCAAGTC TCGCCTCACA ACCACCCCAA ACTGCAGCAA CTGTGGCTGA 540
 AGGCGCATTA CGTGGAGGCC GAGAAGCTGC GCGGCCGACC CCTGGGCGCC GTGGGCAAA 600
 ATCGGGTGC CCGAAAATTT CCATGCGCGC GCACCATCTG GGACGGCGAG GAGACCACT 660
 ACTGCTTCAA GGAGAAGTCG AGGGGTGTCC TGCGGGAGTG GTACGCGCAC AATCCCTACC 720
 CATCGCCGCG TGAGAAGCGG GAGCTGGCCG AGGCCACCGG CCTCACCACC ACCCAGGTCA 780
 GCAACTGGTT TAAGAACCAG AGGCAAGAG ACCGGGCGCG GGAGGCGAAG GAAAGGGAGA 840
 ACACCGAAAA CAATTAATCC TCCTCCAACA AGCAGAACCA ACTCTCTCTC CTGGAAGGGG 900
 GCAAGCCGCT CATGTCCAGC TCAGAAGAGG AATTCTCACC TCCCAAAGT CCAGACCAGA 960
 ACTCGGTCTT TCTGCTGCAG GGCAATATGG GCCACGCCAG GAGCTCAAACT TATCTCTTCC 1020
 CGGGCTTAAC AGCCTCGCAG CCCAGTCACG GCTGCGAGAC CCACCAAGCAT CAGCTCCAAG 1080
 ACTCTCTGCT CGGCCCTCTC ACCTCCAGTC TGGTGGACTT GGGGTCTTAA GTGGGGAGGG 1140
 ACTGGGGCCT CGAAGGGATT CCTGGAGCAG CAACCACTGC AGCGACTAGG GACACTTGTA 1200
 AATAGAAATC AGGAACATTT TTGCAAGCTTG TTTCTGGAGT TGTTTGGCCA TAAAGGAATG 1260
 GTGGACTTTC ACAAATATCT TTTTAAAAAT CAAAACCAAC AGCGATCTCA AGCTTAATCT 1320
 CCTCTCTCTC CCAACTCTTT CCACCTTTTC ATTTTCTTTC CCAATGCAGA GATCAGGG

SEQ ID NO:108 PEW3 Protein sequence

Protein Accession #: NP_005973

1 11 21 31 41 51
 MSMLPSFGFT QEQVACVCEV LQQGGNLERL GRFLWSLPAC DHLHKNESVL KAKAVVAFHR 60
 GNFRELYKIL ESHQSPHNH PKLQLWLKA HYVEAEKLRG RPLGAVGKYR VRRKFLPRT 120
 IWDGEETSYC FKEKSRGVLR EWWAHNPYP PREKRELAEA TGLTTTQVSN WFKNRRQRDR 180
 AAEAKERENT ENNNSSSNKQ NQLSPLEGGK PLMSSSEEFV SPPQSPDQNS VLLQLQNMGH 240
 ARSSNYSLEP LTASQPSHGL QTHQHLQDS LLGLTSSSLV DLGS

SEQ ID NO:109 PFJ8 DNA SEQUENCE

Nucleic Acid Accession #: NM_005069

Coding sequence: 57-2060 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 GGGGCTCCGC GGGCCTGGAG CACGGCCGGG TCTAATATGC CCGGAGCCGA GGC GCGATGA 60
 5 AGGAGAAGTC CAAGAATGCG GCCAAGACCA GGAGGGAGAA GGAAAATGGC GAGTTTACG 120
 AGCTTGCCAA GCTGCTCCCG CTGCCGTCGG CCATCACTTC GCAGCTGGAC AAAGCGTCCA 180
 TCATCCGCCT CACCACGAGC TACCTGAAGA TCGCGCGCGT CTTCCCGGAA GGTTTAGGAG 240
 ACGCGTGGGG ACAGCCGAGC CGCGCCGGGG CCTTGGACGG CGTCGCCAAG GAGCTGGGAT 300
 CGCACTTGCT GCAGACTTTG GATGGATTG TTTTGTGGT AGCATCTGAT GGCAAAATCA 360
 10 TGTATATATC CGAGACCGCT TCTGTCCATT TAGGCCTATC CCAGGTGGAG CTCACGGGCA 420
 ACAGTATTTA TGAATACATC CATCTTCTG ACCACGATGA GATGACCGCT GTCCTACGG 480
 CCCACGACC GCTGCACCAC CACCTGCTCC AAGAGTATGA GATAGAGAGG TCGTTCTTTC 540
 TTCGAATGAA ATGTGTCTTG GCGAAAAGGA ACGCGGGCCT GACCTGCAGC GGATACAAGG 600
 TCATCCACTG CAGTGGCTAC TTGAAGATCA GGCAGTATAT GCTGGACATG TCCCTGTACG 660
 15 ACTCCTGCTA CCAGATTGTG GGGCTGGTGG CCGTGGGGCA GTCGCTGCCA CCCAGTGCCA 720
 TCACCGAGAT CAAGCTGTAC AGTACATGT TCATGTTTCA GGCACGCCCT GACCTGAAGC 780
 TGATATTCCT GGATTCAGG GTGACCGAGG TGACGGGTTA CGAGCCGAG GACCTGATCG 840
 AGAAGACCCT ATACCATCAC GTGCACGGCT GCGACGTGTT CCACCTCCGC TACGCACACC 900
 ACCCTCTGTT GTGAAGGGC CAGGTACCA CCAAGTACTA CCGGCTGCTG TCCAAGCGGG 960
 GCGGGTGGGT GTGGGTGAC AGCTACGCCA CCGTGGTGCA CAACAGCCGC TCGTCCCGGC 1020
 20 CCCACTGCAT GTGAGTGTG AATTATGTAC TCACGGAGAT TGAATACAAG GAACTTCAGC 1080
 TGTCCCTGGA GCAGGTGTCC ACTGCCAAGT CCCAGGACTC CTGGAGGACC GCCTTGTCTA 1140
 CCTCACAAGA AACTAGGAAA TTAGTGAAC CCAAAAATAC CAAGATGAAG ACAAAGCTGA 1200
 GAACAAAACC TTACCCCCCA CAGCAATACA GCTCGTTCCA AATGGACAAA CTGGAATGCG 1260
 GCCAGCTCGG AAAGTGGAGA GCCAGTCCCC CTGCAAGCGC TGCTGCTCCT CCAGAAGTGC 1320
 25 AGCCCCACTC AGAAAGCAGT GACCTTCTGT ACACGCCATC CTACAGCCTG CCCTTCTCCT 1380
 ACCATTACGG ACCTTCTCCT CTGGACTCTC ACGTCTTCAG CAGCAAAAAG CCAATGTTGC 1440
 CGGCCAAGTT CGGGCAGCCC CAAGGATCCC CTTGTGAGGT GGCACGCTTT TTCCTGAGCA 1500
 CACTGCCAGC CACCTCCCGT TGGCAGTGGC ATTATGCCA CCCCCTAGTG CCTAGCAGCT 1560
 30 CGTCTCCAGC TAAAAATCCT CCAGAGCCAC CGGCGAACAC TGCTAGGCAC AGCCTGGTGC 1620
 CAAGCTACGA AGCGCCCGCC GCGCGCGTGC GCAGGTTCGG CGAGGACACC GCGCCCGCGA 1680
 GCTTCCCGAG CTGCGGCCAC TACCGCGAGG AGCCCGCGCT GGGCCCGGCC AAAGCCGCCC 1740
 GCCAGGCCGC CCGGGACGGG GCGCGGCTGG CGCTGGCCCG CGCGGCACC GAGTGCTGCG 1800
 CGCCCCGAC CCCCAGGGCC CCGGGCGCGC CGGCGCAGCT GCCCTTCGTG CTGCTCAACT 1860
 35 ACCACCGGCT GCTGGCCCGG CCGGACCGC TGGGGGGCGC CGCACCCGCC GCCTCCGGCC 1920
 TGGCCTGCGC TCCCGGGCGG CCGGAGCGCG CGACCGGCGC GCTGCGGCTC CGGCACCCGA 1980
 GCGCCGCGC CACCTCCCGC CCGGCGCGC CCTGCGGCA CTACCTGGG GCTCGGTCA 2040
 TCATCACCAA CGGGAGGTGA CCGCTGGCC CCGCGGCCA GGAGCCTGGA CCGGCGCTCC 2100
 CGGGGCTGCG GCGCCACCGA CGCCGGCAAA TGCGCACGAC CTACATTAAT TTATGCAGAG 2160
 40 ACAGCTGTTT GAATTGGACC CCGCGGCCGA CTTGCGGATT TCCACCGCG AGGCCCGCG 2220
 CGCCGTGCC GAGGGCCGAG GAGCGCCCG GTCCGGGCG GTGACCGCCC GCCTCTGTCC 2280
 TGCGAGGGCC GGTGCGACCC AGTTGCTGGG GGCTTGGTTT CCTACCTTG AAATCGGGCT 2340
 TCACGCTGCT TGCCTTGTC CCAACGTTCC ACAACAGTCC CGCTGGGGGA TTGAAGCGGT 2400
 TTCATCCGC AAATATCCTC CACTTTCAGG AGGGAAGACC CACCTACCA CAGTCCGCTC 2460
 45 TTCCAAGTGC ACGCGAGACC TGGGAGGGGA CGCTGTGTC ACGAGCCCTT TTAGATGCTT 2520
 AGGTGAAGGC AGAAGTGATG ATTGTAAGTC CCATGAATAC ACAACTCCAC TGTCTTTAAA 2580
 AGTCATTCAA GAGTCTCAT ATTTTGTGTT TTATTTAACC CTTTCTTCAA TACAAAAAGC 2640
 CAACAAACCA AGACTAAGGG GTGACCATG CAATTCCATT TTGTGCTGT GAACATAGGT 2700
 GTGCTTCCA AATACATTAA CAAGCTCTTA CTCCCCCTA ACCCTATGA ACTCTTGATA 2760
 50 ACACCAAGAG TAGCACCTTC AGAATATATT GAATAGGCAT TAAATGCAAA AATATATATG 2820
 TAGCCAGACA GTTATGAGA ATGACCTGT CAAGCTTCAT TATTACGTGG CAAAATCCCT 2880
 CTGGGCCACA CAGATCTGTA ATTCACTAGG CTCGTGTTG CTACAAATAG TGCTAATAAA 2940
 GTTAAATTGC ACGTGCAATA CGGAACACTG TCAATGGACT GCACCTGTG AAGGAAAAAC 3000
 ATGCTTAAGG GGGTGTAAATG AAAATGATGT AGACATTTTA AGCATTTTCT ACACAGCGAG 3060
 55 AAAACTTCGT AAGAACATGT TACGTGTGCA ACAGGTAAAC AGAAATCCTT TCATAAAGCA 3120
 CCAGCAGTGT TAAAAAATG AGCTTCCATT AATTTTACT TTTTATGGG TTTGCTTAAA 3180
 GATCTCAACA TGAAAAATC CTGTCTGCG TCTGAACGTC ACAATGCATT GAACCGCCGT 3240
 CCTTCAATTT TCTTCACT ATCAACACTG CAGCATTTTG CTGCTTATC AAAATGGTTT 3300
 ATTTTAGGAA ACTTTTCCA CTTTCTGAA TGGAAAGAGG TTTTCAAAA TGTTTTAAAC 3360
 60 TCATCGTTCT AAAATCAAGT GCACCTACAC CAACTGCTCT CAAAATGTGA ACTGACTTTT 3420
 TTTTTTTTTT TTTTGCAAC CCGTGTGTC TTAGTGAGGA CCGACACAA TCCCTACAGG 3480
 GTGCTGTGTA GTGGGCTCA TGGTAAGAGT CACAATTGCA AAATTAGGA CCGTGGGTCA 3540
 TGCAGCGAAG GGGCTGGATG GTAGGAAGGG ATGTGCCCGC CTCCTCACGC ACTCAGCTAT 3600
 65 ACCTATTCA CAGCTCCTTG TGAGTGTGTG CACAGGAAAT AAGCCGAGGG TATTATTTT 3660
 TTATGTTTAT GAGTCTGTA ATTAACCGT GATTCTTGAA AGGTGTAGGT TTGATTACTA 3720
 GGAGATACCA CCGACATTTT TCAATAAAGT ACTGCAAAAT GCTTTTGTG CTACCTTGTT 3780
 ATTAACTTT GGGGCTGTAT TTAGTAAAAA TAAATCAAGG CTATCGGAGC AGTTCAATAA 3840
 CAAAGTTTAC TGTGAGAAA AAAGACCCTA TCATAGATTT ACAAG

SEQ ID NO:110 PFJ8 Protein sequence:
 Protein Accession #: NP_005060.1

1 11 21 31 41 51
 MKEKSKNAAK TRREKENGFE YELAKLLPLP SAITSQLDKA SHRLTTSYL KMRAVFPEGL 60
 75 GDAWGQPSRA GPLDGVAKEL GSHLLQTLTG FVFVVASDGG IMYISETASV HLGLSQVELT 120
 GNSIYIYIHP SDHDEMTAVL TAHQPLHHHL LQYEIERSF FLRMKCVLAK RNAGLTCSGY 180
 KVIHCSGYLK IRQYMLDMSL YDSYQIVGL VAVGQSLPPS AITEIKLYSN MFMFRASLDL 240

KLIFLDSRVTV ETVGYEPQDL IEKTLYHHVH GCDVFHLRYA HLLLLVKQGV TTKYYRLLSK 300
 RGGWVWVQSY ATVVHNSRSS RPHCIVSVNY VLTEIEYKEL QLSLEQVSTA KSQDSWRTAL 360
 STSQETRLKV KPNKNTMKTK LRTNPYPFQQ YSSFQMDKLE CGQLGNWRAS PPASAAAPPE 420
 LQPHSESDL LYTPSYSLPF SYHYGHFLD SHVFSSKKPM LPAKFGQPOG SPCEVARFFL 480
 STLPAEGECQ WHYANPLVPS SSSPAKNPPE PPANTARHSL VPSYEAPAAA VRRFGEDTAP 540
 PSFPCSGHYR EEPALGPAKA ARQAARDGAR LALARAAPEC CAPPTPEAPG APAQLPFVLL 600
 NYHRVLARRG PLGGAAPAAS GLACAPGGPE AATGALRLRH PSPAATSPPG APLPHYL GAS 660
 VIITNGR

SEQ ID NO:111 PFJ7 DNA SEQUENCE

Nucleic Acid Accession #: NM_006549

Coding sequence: 1-1254 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 ATGAACGGAC GCTGCATCTG CCCGTCCTG CCCTACTCAC CCGTCAGCTC CCCGAGTCC 60
 TCGCCTCGGC TGCCCGGCG GCGACAGTG GAGTCTCACC ACGTCTCCAT CACGGGTATG 120
 CAGGACTGTG TGCAGCTGAA TCAGTATACC CTGAAGGATG AAATTGGAAA GGGCTCCTAT 180
 GGTGTGCTCA AGTIGGCCCTA CAATGAAAAAT GACAATACCT ACTATGCAAT GAAGGTGCTG 240
 TCCAAAAAGA AGCTGATCCG GCAGGCCGGC TTTCCACGTC GCCCTCCACC CCGAGGCACC 300
 CGGCCAGCTC CTGGAGGCTG CATCCAGCCC AGGGGCCCA TTGAGCAGGT GTACCAGGAA 360
 ATTGCCATCC TCAAGAAGCT GGACCACCCC AATGTGGTGA AGCTGGTGA GGTCTGGAT 420
 GACCCCAATG AGGACCATCT GTACATGGTG TTCGAAGTGG TCAACCAAGG GCCCGTGATG 480
 GAAGTGCCCA CCTCAAACC ACTCTCTGAA GACCAGGCC GTTTCTACTT CCAGGATCTG 540
 ATCAAAGGCA TCGAGTACTT ACATACCAG AAGATCATCC ACCGTGACAT CAAACCTTCC 600
 AACCTCTGG TCGGAGAAGA TGGGCACATC AAGATCGCTG ACTTTGGTGT GAGCAATGAA 660
 TTCAAGGGCA GTGACGCGCT CCTCTCCAAC ACCGTGGGCA CGCCCGCCTT CATGGCAGCC 720
 GAGTCGCTCT CTGAGACCCG CAAGATCTTC TCTGGGAAGG CTTGGATGT TTGGGCCATG 780
 GGTGTGACAC TATACTGCTT TGCTTTGGC CAGTGCCCAT TCATGGACGA GCGGATCATG 840
 TGTTTACACA GTAAGATCAA GAGTCAGGCC CTGGAATTTC CAGACCAGCC CGACATAGCT 900
 GAGGACTTGA AGGACCTGAT CACCCGTATG CTGGACAAGA ACCCCGAGTC GAGGATCGTG 960
 GTGCCGAAAA TCAAGCTGCA CCCCTGGGTC ACGAGGCATG GGGCGGAGCC GTTGCCGTCG 1020
 GAGGATGAGA ACTGCACGCT GGTGGAAGTG ACTGAAGAGG AGGTCGAGAA CTCAGTCAAA 1080
 CACATCCCCA CTTGGCAAC CGTGATCCTG GTGAAGACCA TGATACGTAA ACGCTCCTTT 1140
 GGGAAACCAT TCGAGGGCAG CCGGCGGGAG GAACGCTCAC TGTCAGCGCC TGGAACTTG 1200
 CTCACCAAAA AACCAACCA GGAATGTGAG TCCCTGTCTG AGCTCAAGAC CTAGAAAAATA 1260
 AGTCCCCTTC CTGCCTGTTG CAAAGTAACG TAAGAGTTCC CTCACCCGAG TGGATGCAGA 1320
 CGTCTTGCT GTCAGCCACC TTCTTCATA CACATAGCCA GCCCAGGGTG ACCAGAACGT 1380
 CCCAGGACAG ATGAGGCTTT GTGTCCTTAT GAGAGTGGA GAACCTGGTG GGCACCCCTG 1440
 GTGCAGGTGC TGTGGTGGGT GGGGACCCCA CTGCCTTTCC CACTGAGCAC ATCATGGCTA 1500
 CCTGACTTGG TGGGAGTTCC ATTCAGTCAC TTCTGTTTCT TAAACATAGC TTTACTGAGG 1560
 TACAATTAC ATACCATGTA ATTACCCAC GGAAGTGTGA TGATTCAGTG GTTTCTAATA 1620
 CACACTTCTG CAGCCATTAC CACCGTCAAC TTTACGACAT TTTCATCAGC CCAAGAAGAC 1680
 ACCCTACAT CTTAGCTGT CCCCATCCAA CTCCCCACC CCAGTAACCA CTCAGAATAG 1740
 GTATGGATTT GCCTATTCTG GACGTTTCGT ATAAATGGCG TCATACACTA AAAAAAAAAA 1800
 AAAA

SEQ ID NO:112 PFJ7 Protein sequence:

Protein Accession #: NP_006540.1

1 11 21 31 41 51
 | | | | |
 MNGRCICPSL PYSVSSPQS SPRLPRRPTV ESHHVSITGM QDCVQLNQYT LKDEIGKGSY 60
 GVVKLAYNEN DNTYYAMKVL SKKKLIRQAG FPRRPPRGT RPAPGGCIQF RGPQVYQF 120
 IAILKKLDHP NVVKLVEVLD DPNEDHLYMV FELVNQGPVM EVPTLKPLSE DQARFYFDL 180
 IKGIEYLYHQ KIIHRDIKPS NLLVGEDGHI KIADFGVSNE FKGS DALLSN TVGTPAFMAP 240
 ESLSETRKIF SGKALDVWAM GVTLYCFVFG QCFPMDERIM CLHSKIKSQA LEFPDQPDIA 300
 EDLKLITRM LDKNPESRIV VPEIKLHPWV TRHGAELPS EDENCTLVEV TEEEVENSVK 360
 HIPSLATVIL VKTMRKRSF GNPFEGRRE ERSLSAPGNL LTKKPTRECE SLSELKT

SEQ ID NO:113 PFJ6 DNA SEQUENCE

Nucleic Acid Accession #: NM_021810

Coding sequence: 1-429 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 ATGAAACCTC TGATATGGAC ATGGTCAGAT GTTGAAGGCC AGAGGCCGGC TCTGCTCATC 60
 TGCACAGCTG CAGCAGGACC CACGCAGGGA GTTAAGGGTT ATGGCAAGCC CTTTGAGCCA 120
 AGAAGTGTGA AAAACATACA CTCTACTCCT GCTTACCCAG ATGCCACAAT GCACAGACAA 180
 CTCCTGGCTC CGGTGGGAAG AGGATGGGA GAGACATTGA ATCAGAACT CCATGTTGCC 240
 AATGTGCTGG AAGATGACCC CGGCTACCTA CCTCACGTCT ACAGCGAGGA AGGGGAGTGT 300
 GGAGGGGCCC CATCCCTCAG CTCTCTGGCC AGCTTGAAC AGGAGTTGCA ACCTGATTG 360

CTGGACTCTT TGGGTTCAAA AGCGACTCCG TTTGAGGAAA TATATTCAGA GTCAGGTGTT 420
 CCTTCCTAA

SEQ ID NO:114 PFJ6 Protein sequence:
 Protein Accession #: NP_068582.1

1 11 21 31 41 51
 MKPLIWTWSD VEGQRPALLI CTAAAGPTQG VKGYGKPFEP RSVKNIHSTP AYPDATMHRQ 60
 LLAPVEGRMA ETLNQKLHVA NVLEDDPGYL PHVYSEEGEC GGAPSLSSLA SLEQELQPD 120
 LDSLGSKATP FEEIYSESGV PS

SEQ ID NO:115 PFJ5 DNA SEQUENCE

Nucleic Acid Accession #: NM_006361
 Coding sequence: 131-985 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 CGAATGCAGG CGACTTGCGA GCTGGGAGCG ATTAAAAACG CTTTGGATTG CCCCAGGCTG 60
 GGTGGGGAGA GCGAGCTGGG TGCCCCCTAG ATTCCCCGCC CCCGCACCTC ATGAGCCGAC 120
 CCTCGGCTCC ATGGAGCCCC GCAATTATGC CACCTTGGAT GGAGCCAAGG ATATCGAAGG 180
 CTTGCTGGGA GCGGGAGGGG GCGGGAATCT GGTGCGCCAC TCCCCTCTGA CCAGCCACCC 240
 AGCGGCGCCT ACGCTGATGC CTGCTGTCAA CTATGCCCCC TTGGATCTGC CAGGCTCGGC 300
 GGAGCCGCCA AAGCAATGCC ACCCATGCCC TGGGGTGCCC CAGGGGACGT CCCAGCTCC 360
 CGTGCCCTTAT GGTTACTTTG GAGGCGGGTA CTACTCCTGC CGAGTGTCCC GGAGCTCGCT 420
 GAAACCTGT GCCCAGGCAG CCACCTTGGC CGCGTACCCC GCGGAGACTC CCACGGCCGG 480
 GGAAGAGTAC CCCAGTCGCC CCACTGAGTT TGCCTTCTAT CCGGGATATC CGGGAACCTA 540
 CCACGCTATG GCCAGTTACC TGGACGTGTC TGTGGTGCA GACTCTGGGTG CTCCTGGAGA 600
 ACCGCGACAT GACTCCCTGT TGCTGTGGA CAGTTACCAG TCTTGGGCTC TCGCTGGTGG 660
 CTGGAACAGC CAGATGTGTT GCCAGGGAGA ACAGAACCCA CCAGTCCCT TTTGGAAGGC 720
 AGCATTTGCA GACTCCAGCG GGCAGCACCC TCCTGACGCC TCGCCTTTC GTCGCGGCCG 780
 CAAGAAACGC ATTCCGTACA GCAAGGGGCA GTTGCGGGAG CTGGAGCGGG AGTATGCGGC 840
 TAACAAGTTC ATCACCAAGG ACAAGAGGCG CAAGATCTCG GCAGCCACCA GCCTCTCGGA 900
 GCGCCAGATT ACCATCTGGT TTCAGAACCG CCGGGTCAAA GAGAAGAAGG TTCTCGCCAA 960
 GGTGAAGAAC AGCGTACCC CTAAAGAGAT CTCCTTGCC TGGTGGGAGG AGCGAAAGTG 1020
 GGGGTGTCCT GGGGAGACCA GAAACCTGCC AAGCCCAGGC TGGGGCCAAG GACTCTGCTG 1080
 AGAGGCCCCC AGAGACAACA CCCTTCCCAG GCCACTGGCT GCTGGACTGT TCCTCAGGAG 1140
 CGGCTGGGT ACCCAGTATG TGCAGGGAGA CGGAACCCA TGTGACAGGC CCACTCCACC 1200
 AGGGTTCCCA AAGAACCTGG CCCAGTCATA ATCATTATC CTCACAGTGG CAATAATCAC 1260
 GATAACCACT

SEQ ID NO:116 PFJ5 Protein sequence:
 Protein Accession #: NP_006352.1

1 11 21 31 41 51
 MEPMNYATLD GAKDIEGLLG AGGGRNLVAH SPLTSHPAAP TLMPAVNYAP LDLPGSAEPP 60
 KQCHPCPGVP QGTSPAPVPY GFYGGGYSC RVSRSCLKPC AQAATLAAYP AETPTAGEEY 120
 PSRPTEFAFY PGYPGTYHAM ASYLDVSVVQ TLGAPGEPRH DSIPLVDSYQ SWALAGGWNS 180
 QMCCQGEQNP PGPFWKAAFA DSSGQHPPDA CAFRRGRKKR IPYSKGQLRE LEREYAANKF 240
 ITKDKRRKIS AATSLSERQI TIWFQNRVRK EKKVLAKVKN SATP

SEQ ID NO:117 PFJ4 DNA SEQUENCE

Nucleic Acid Accession #: NM_005628
 Coding sequence: 591-2216 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 GTAACCGCTA CTCCCGGACA CCAGACCACC GCCTTCCGTA CACAGGGGCC CGCATCCAC 60
 CCTCCCGGAC CTAAGAGCCT GGGTCCCTTG TTCCGGAGG TCCGCTTCCC GGCCCCCAGA 120
 TTTGGGCATC CCAGCCCTCA GTGTCCAAGA CCCAGGCAGC CCGGGTCCCC GCCTCCCGGA 180
 TCCAGGCGTC CGGGATCTGC GCCACCAGAA CCTAGCCTCC TGCAGACCTC CGCATCTGG 240
 GGGCACTCAA CCTCTGGAG CCAAGGGGCC CAGTCCAC CCAGAGAAAC TCTCGTATTC 300
 CCAGCTCCTA GGGCCAAAGG ACCCGGGCGC TCCGAATCC CAGCTTTCGG ACATCTGGCA 360
 CACGGGGCAG AGCAGAGAAG CTCAGCGCCC AGCCTGGGGA ATTAAACAC TCCAGCTTCC 420
 AAGAGCCAAG GAACTTCACT GCTGTGAAC CACAACCTA AGGAGCCCTC CAAAGTTCCA 480
 GTCTCCAGGT GCTGTACTC AACTCAGTCC TAGGAACGTC GGGTCTGGG AAGGAGCCCA 540
 AGCGCTCCCA GCCAGTTC AGGCGCTAAG AAACCCCGGT GCTTCCCATC ATGGTGGCCG 600
 ATCTCCTCG AGACTCAAAG GGGCTCGCAG CGGCGGAGCC CACCGCCAAC GGGGGCCTGG 660
 CGTGGCCTC CATCGAGGAC CAAGGCGCGG CAGCAGGCGG CTAAGTGGT TCCCGGACC 720
 AGGTGCGCCG CTGCCTTCA GCGAACCTGC TTGTGCTGCT GACAGTGGTG GCCGTGGTGG 780
 CCGCGGTGGC GCTGGGACTG GGGGTGTCGG GGGCCGGGGG TCGCTGGCG TTGGCCCCGG 840

AGCGCTTGAG CGCCTTCGTC TTCCCGGGCG AGCTGCTGCT GCGTCTGCTG CGGATGATCA 900
 TCTTGCCGCT GGTGGTGTGC AGCTTGATCG GCGGCGCCGC CAGCCTGGAC CCCGGCGCGC 960
 TCGGCCGTCT GGGCGCCTGG GCGCTGCTCT TTTTCTGGT CACCACGCTG CTGGCGTCGG 1020
 CGCTCGGAGT GGGCTTGGCG CTGGCTCTGC AGCCGGGGCG CGCCTCCGCC GCCATCAACG 1080
 CCTCCGTGGG AGCCGCGGGC AGTGCCGAAA ATGCCCCAG CAAGGAGGTG CTCGATTCTG 1140
 TCCTGGATCT TGCAGAAAT ATCTCCCTT CCAACCTGGT GTCAGCAGCC TTTCGCTCAT 1200
 ACTCTACCAC CTATGAAGAG AGGAATATCA CCGGAACCAG GGTGAAGGTG CCCGTGGGGC 1260
 AGGAGGTGGA GGGATGAAC ATCTGGGCT TGGTAGTGT TTGGTGTGG 1320
 CGCTGCGGAA GCTGGGGCCT GAAGGGGAGC TGCTTATCCG CTCTTCAAC TCCTTCAATG 1380
 AGGCCACCAT GGTCTGTGTC TCCTGGATCA TGTGGTACGC CCTGTGGGC ATCATGTTCC 1440
 TGGTGGCTGG CAAGATCGTG GAGATGGAGG ATGTGGGTTT ACTCTTGCC CGCCTTGGCA 1500
 AGTACATTCT GTGCTGCTG CTGGGTACG CCATCCATGG GCTCCTGGTA CTGCCCCCTCA 1560
 TCTACTTCT CTTCACCCGC AAAAACCCCT ACCGCTTCT GTGGGGCATC GTGACGCCGC 1620
 TGGCCACTGC CTTTGGGACC TCTTCCAGTT CCGCCACGCT GCGCTGATG ATGAAGTGCG 1680
 TGGAGGAGAA TAATGGCGTG GCCAAGCACA TCAGCCGTTT CATCCTGCC ATCGGCGCCA 1740
 CCGTCAACAT GGACGGTGCC GCGCTCTTC AGTGGCTGGC CGCAGTGTTC ATTGCACAGC 1800
 TCAGCCAGCA GTCCTTGGAC TTCGTAAAGA TCATCACCAT CTTGGTACG GCCACAGCGT 1860
 CCAGCGTGGG GGCAGCGGGC ATCCCTGCTG GAGGTGTCT CACTCTGGCC ATCATCTCG 1920
 AAGCAGTCAA CTTCCCGTC GACCATATCT CCTTGATCT GGCTGTGGAC TGGCTAGTCG 1980
 ACCGGTCTTG TACCTGCTCT AATGTAGAAG GTGACGCTCT GGGGGCAGGA CTCCTCAAA 2040
 ATTAGTGTGA CCGTACGGAG TCAGAGAAGCA CAGAGCCTGA GTTGATACAA GTGAAGAGTG 2100
 AGCTGCCCTT GGATCCGCTG CCACTCCCA CTGAGGAAGG AAACCCCTC CTCAAAACT 2160
 ATCGGGGGCC CGCAGGGGAT GCCACGGTCG CCTCTGAGAA GGAATCAGTC ATGTAAACCC 2220
 CGGGAGGGAC TTCCCTGCC CTGCTGGGGG TGCTCTTGG ACCTGGATT ATGAGGAATG 2280
 GATAAATGGA TTAGCTAGGG CTCTGGGGT CTGCTGCAC ACTCTGGGA GCCAGGGGCC 2340
 CCAGCACCTT CCAGGACAGG AGATCTGGGA TGCTGGCTG CTGGAGTACA TGTGTTACA 2400
 AGGGTACTC CTCAAAACCC CAGTTCTCA CTCATGTCCC CAACTCAAGG CTAGAAAACA 2460
 GCAAGATGGA GAAATAATGT TCTGCTGCGT CCCACCGTG ACCTGCCTGG CCTCCCTGT 2520
 CTCAGGGAGC AGGTACACAG TCACCATGGG GAATTCTAGC CCCCCTGGG GGGATGTTAC 2580
 AACACCATGC TGGTTATTTT GCGGCTGTA GTTGTGGGG GATGTGTGTG TGCAGTGTG 2640
 TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG TTCTGTGACC TCCTGTCCC ATGGTACGTC 2700
 CCACCTGTC CCCAGATCCC CTATCCCTC CACAATAACA GAAACACTCC CAGGGACTCT 2760
 GGGGAGAGGC TGAGGACAAA TACCTGCTGT CACTCCAGAG GACATTTTTT TTAGCAATAA 2820
 AATTGAGTGT CAACTATTTA AAAAAAAAAA AAAAAA

SEQ ID NO:118 PFJ4 Protein sequence:

Protein Accession #: NP_005619.1

1 11 21 31 41 51
 | | | | |
 MVADPPRDSK GLAAAEPTAN GGLALASIED QGAAAGGYCG SRDQVRRCLR ANLLVLLTVV 60
 AVVAGVALGL GVSAGGALA LGPERLSAFV FPGELLRL RLMIPLVVC SLIGGAASLD 120
 PGALGRLGAW ALLFLVTL LASALGVGLA LALQPGAASA AINASVGAAG SAENAPSKEV 180
 LDSFLDLARN IFFSNLVSA FRSYSTYEE RNITGTRVKV PVGQEVGMN ILGLVVFAIV 240
 FGVALRLKLP EGELLIRFFN SFNEATMVLV SWIMWYAPVG IMFLVAGKIV EMEDVGLLEA 300
 RLGYKILCL LGHAIHLLV LPLIYFLFR KNPYRFLWGI VTPLATAFGT SSSSATLPLM 360
 MKCVEENNGV AKHISFILP IGATVNMDGA ALFQCVAAVF IAQLSQSLD FVKIITLVT 420
 ATASSVGAAG IPAGGVLTIL ILEAVNLPV DHISLILAVD WLVDRSCTVL NVEGDALGAG 480
 LLQNYVDRTE SRSTEPELIQ VKSELPLDPL PVPTEGNPL LKHYRGPAGD ATVASEKESV 540
 M

SEQ ID NO:119 PFJ3 DNA SEQUENCE

Nucleic Acid Accession #: NM_006708

Coding sequence: 88-642

(underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 CTAGTTAAGG CGGCACAGGG CCGAGGCGTA GTGTGGGTGA CTCCTCCGTT CCTTGGGTCC 60
 CGTCGTCTGT GATACTGCAG TTCAGCCATG GCAGAACCGC AGCCCCCGTC CGGCGGCCTC 120
 ACGGACGAGG CCGCCCTCAG TTGCTGCTCC GACGCGGACC CCAGTACCAA GGATTTTCTA 180
 TTGCAGCAGA CCATGTACG AGTGAAGGAT CTAAGAAGT CACTGGATTT TTAATACTAGA 240
 GTTCTTGAA TGACGCTAAT CCAAAAATGT GATTTTCCCA TTATGAAGTT TTAATCTAC 300
 TTCTTGGCTT ATGAGGATAA AAATGACATC CCTAAAGAAA AAGATGAAAA AATAGCCTGG 360
 GCGCTCTCCA GAAAAGCTAC ACTTGAGCTG ACACACAATT GGGGCACTGA AGATGATGCG 420
 ACCCAGAGTT ACCACAATGG CAATTCAGAC CCTCGAGGAT TCGGTATAT TGAATTTGCT 480
 GTTCTGTATG TATACAGTGC TTGTAAGAGG TTGTAAGAAC TGGGAGTCAA ATTTGTGAAG 540
 AAACCTGATG ATGGTAAAT GAAAGGCTG GCATTTATC AAGATCTCGA TGGCTACTGG 600
 ATTGAAATTT TGAATCTTAA CAAAATGGCA ACCTTAATGT AGTGCTGTGA GAATTTCTCT 660
 TTGAGATTTT AGAAGAAAGG AAACAATGTG ATTCAAGATA TTACATACC AGAAGCATCT 720
 AGGACTGATG GATCACTGTC CCGATTCAA TTATTCTTCA GTCCATTTC CCTTCCTATT 780
 TCAGCTGTTT CTTTACCT AACTGTTTCA TCATTCTGGT TTCAAGCAG TGCTTTATCT 840
 CATGCTCTTG AATATAGTTG TGTAACTTTA TTTTATAGT AATAATTAGA ACAGTTCCCT 900
 TCAGAGGCTG CATTTGCTT CTCTGCCAC CTAATATTTA CTCCCTTCA AATCTGCTT 960
 TGAATCATCA TTTTAAAAA AAAATTAACA TGTTTTGTG GTAGTTATCT TCTGGGGTTT 1020
 CAATTCCTCA GAAACAATTT TTTTCAAC GGAAGGAAA GAACACTAGT GTTCTTTCAG 1080
 TAAAGTACAA AGTGTTTTATT TTACAAAAGA GTAGGTACTC TTGAGAGCAA TTCAATCAT 1140

5
10
15

GCTGACAAGG ATACTGATAG AAAAAGTGAT TTCTTCTTAT TATAAAGTAC ATTTAAAGTT 1200
CAAGGACTAA CCTTATTTAT TTGGGAAAGG GGAGGAGGAA GGAAATGATA TGGTACCCAG 1260
ACACTGGGCT AGGCTGCAAC TTTATCTCAT TTAATACTCC CAGCTGTCAT GTGAGAAAGA 1320
AAGCAGGCTA GGCATGTGAA ATCACTTTCA TGGATTATTA ATGGATTAA GAGGGCATCA 1380
ATCAGCTCAA CTCAGATTIT CATAATCATT TTAGTATTT AGATTGTGCC TCAAAGTTGT 1440
AGTACCTCAC AATACCTCCA CTGGTTTCCT GTTGTAAGAAA CCTTCAGTGA GTTTGACCAT 1500
TGTGCTCTTG GCTCTTGGGC TGGAGTACCG TGGTGAGGGA GTAAACACTA GAAGTCTTTA 1560
GTACAAAAC TCTTAGGGA CACCTGGTGA TTCCTACACA AGTGATGTTT ATATTTCTCA 1620
TAAAGAGTCT TCCCTATCCC AAGGCTTCA TGATGCCAGT AGCCATATAT GATAAATTAT 1680
GTTTCAGTAT AACTTAGTTA TCAGAAATCA GCTCAGTGGT CTCCCCGCC ATGATTACA 1740
TTTGTAGAGT TTTTAAAAAT CAAAGTGATT TTGAAAATCT CTAATGGCTC AGAAAATAAA 1800
AACATCCAGT TTGTGGATGA CTATATTTAG ATTTCTCTAG ACTCTAGTGG AAGACCTTTG 1860
GAAAGGCCAT GCCAACCGTG CTGTACTGCT TAGAAGCACT TTATGTTTCC TTTTGGGTG 1920
AAATGGATTT ATGTGAGTGC TTTAAACAAA TAGCAATACT TATAGACTGA AATAAAATGA 1980
AACTCAAAAT AAG

SEQ ID NO:120 PFJ3 Protein sequence:

Protein Accession #: NP_006699.1

20
25

1 11 21 31 41 51
MAEPQPPSGG LTDEAALSCC SDADPSTKDF LLQQTMLRVK DPKKSLDFYT RVLGMTLIQK 60
CDFFIMKFSL YFLAYEDKND IPKEKDEKIA WALSRKATLE LTHNWGTEDD ATQSYHNGNS 120
DPRGFHIGI AVPDVYSACK RFEELGVKFFV KKPDDGKMKMG LAFIQDPDGY WIEILNPNKM 180
ATLM

SEQ ID NO:121 PFJ2 DNA SEQUENCE

Nucleic Acid Accession #: NM_002867

Coding sequence: 70-729 (underlined sequences correspond to start and stop codons)

30
35
10
15
20

1 11 21 31 41 51
CCGACGCCAG GTCCTGCCGT CCCGCCGACC GTCGGGGAGC GAACCCGTCG TCCCGCACTG 60
GAGTCCGCGA TGGCTTCAGT GACAGATGGT AAACATGGAG TCAAAGATGC CTCTGACCAG 120
AATTTGACT ACATGTTAA ACTGCTTATC ATTGGCAACA GCAGTGTGG CAAGACCTCC 180
TTCTCTTGC GCTATGCTGA TGACACGTTT ACCCCAGCCT TCGTTAGCAC CGTGGGCATC 240
GACTTCAAGG TGAAGACAGT CTACCGTCAC GAGAAGCGGG TGAAACTGCA GATCTGGGAC 300
ACAGCTGGGC AGGAGCGGTA CCGGACCATC ACAACAGCCT ATTACCGTGG GGCCATGGGC 360
TTCATTCTGA TGTATGACAT CACCAATGAA GAGTCCTTCA ATGCTGTCCA AGACTGGGCT 420
ACTCAGATCA AGACCTACTC CTGGGACAAT GCACAAGTTA TTCTGGTGGG GAACAAGTGT 480
GACATGGAGG AAGAGAGGGT TGTTCCTACT GAGAAGGGCC AGCTCCTTGC AGAGCAGCTT 540
GGGTTTGATT TCTTTGAAGC CAGTGCAAAG GAGAACATCA GTGTAAGGCA GGCCTTTGAG 600
CGCTGTGGT ATGCCATTTG TGACAAGATG TCTGATTTCG TGGACACAGA CCCGTCGATG 660
CTGGGCTCCT CCAAGAACAC GCGTCTCTCG GACACCCAC CGCTGCTGCA GCAGAACTGC 720
TCATGCTAGC AAGGCCACC TTCTGACCT CCCCTCATG TGGCCCCACA CCAAGTCTG 780
CTTCTCCCTG TTACACACTG TCCGCTCT

SEQ ID NO:122 PFJ2 Protein sequence:

Protein Accession #: NP_002858.1

15
20

1 11 21 31 41 51
MASVTDGKHG VKDASDQNFY YMFKLLIGN SSVGKTSFLL RYADDTFIPA FVSTVGIDFK 60
VKTVYRHEKR VKLQIWDTAG QERYRTITTA YYRGAMGFIL MYDITNEESF NAVQDWATQI 120
KTYSWDNAQV ILVGNKCDME EERVVPTEKG QLLAEQLGFD FFEASAKENI SVRQAFERLV 180
DAICDKMSDS LDTDPSMLGS SKNTRLSDTP PLLQNCSC

SEQ ID NO:123 PFJ1 DNA SEQUENCE

Nucleic Acid Accession #: NM_001844

Coding sequence: 158-4621 (underlined sequences correspond to start and stop codons)

5
10

1 11 21 31 41 51
ACGCAGAGCG CTGCTGGGCT GCCGGGTCTC CCGCTTCTC CTCCTGCTCC AAGGGCCTCC 60
TGATGAGGG CGCGGTAGAG ACCCGGACCC GCGCCGTGCT CCGCGCTTT CGCTGCGCTC 120
CGCCCGGGCC CGGCTCAGCC AGGCCCGCG GTGAGCCATG ATTCGCTCG GGGCTCCCCA 180
GTCGCTGGTG CTGCTGACG TGCTGTCGC CGCTGTCCTT CGGTGTCAGG GCCAGGATGT 240
CCAGGAGGCT GGCAGCTGTG TGCAGGATGG GCAGAGGTAT AATGATAAGG ATGTGTGGAA 300
GCCGGAGCCC TGCCGGATCT GTGTCTGTGA CACTGGGACT GTCCTCTGCG ACGACATAAT 360
CTGTGAAGAC GTGAAAGACT GCCTCAGCCC TGAGATCCCC TTCGGAGAGT GCTGCCCAT 420
CTGCCAACT GACCTCGCCA CTGCCAGTGG GCAACCAGGA CCAAAGGGAC AGAAAGGAGA 480
ACCTGGAGAC ATCAAGGATA TTGTAGGACC CAAAGGACCT CCGGGCCTC AGGGACCTGC 540

5 AGGGGAACAA GGACCCAGAG GGGATCGTGG TGACAAAGGT GAAAAAGGTG CCCCTGGACC 600
 TCGTGGCAGA GATGGAGAAC CTGGGACCCC TGGAAATCCT GGCCCCCTG GTCTCCCCGG 660
 CCCCCTGGT CCCCCTGGT TTGGTGGAAA CTTTGTCTCC CAGATGGCTG GAGGATTGA 720
 TGAAAAGGCT GGTGGCGCCC AGTTGGGAGT AATGCAAGGA CCAATGGGCC CCATGGGACC 780
 TCGAGGACCT CCAGGCCCTG CAGGTGCTCC TGGGCCTCAA GGATTCAAG GCAATCCTGG 840
 TGAACCTGGT AAACCTGGT TCTCTGGTCC CATGGTCCC CGTGGTCTC CTGGTCCCC 900
 TGGAAAGCCT GGTGATGATG GTGAAGCTGG AAAACCTGGA AAAGCTGGTG AAAGGGGTCC 960
 GCCTGTCTCT CAGGGTGCTC GTGGTTTCCC AGGAACCCCA GGCCTTCTG GTGTCAAAGG 1020
 10 TCACAGAGGT TATCCAGGCC TGGACGGTGC TAAGGGAGAG GCGGGTGCTC CTGGTGTGAA 1080
 GGGTGAGAGT GGTCCCCCG GTGAGAACGG ATCTCCGGGC CCAATGGGTC CTCGTGGCCT 1140
 GCCTGGTGAA AGAGGACGGA CTGGCCCTGC TGGCGCTGCG GGTGCCCGAG GCAACGATGG 1200
 TCAGCCAGGC CCCGACGGTC CTCCGGGTCC TGTGGTCTCT GCTGGTGGTC CTGGCTTCCC 1260
 TGGTGCTCTT GGAGCCAAAG GTGAAGCCGG CCCCCTGGT GCCCCTGGTC CTGAAGGTGC 1320
 TCAAGGTCTT CGCGGTGAAC CTGGTACTCC TGGGTCCCCT GGGCCTGCTG GTGCCTCCGG 1380
 15 TAACCTGGA ACAGATGGAA TTCTGGAGC CAAAGGATCT GCTGGTGCTC CTGGCATTGC 1440
 TGGTGCTCTT GGCTTCCCTG GGCCACGGGG TCCTCCTGGC CCTCAAGGTG CAACTGGTCC 1500
 TCTGGGCCCG AAAGGTGAGA CGGGTGAACC TGGTATTGCT GGCTTCAAAG GTGAACAAGG 1560
 CCCCAGGGA GAACCTGGCC CTGCTGGCCC CCAGGAGGCC CTGGACCCG CTGTGGAAGA 1620
 AGGCAAGAGA GGTGCCCGTG GAGAGCCTGG TGGCGTTGGG CCCATCGGTC CCCCTGGAGA 1680
 20 AAGAGGTGCT CCCGAAACC GCGGTTTCCC AGGTCAAGAT GGTCTGGCAG GTCCCAAGGG 1740
 AGCCCCGGA GAGCGAGGGC CAGTGGTCT TGCTGGCCCC AAGGGAGCCA ACGGTGACCC 1800
 TGGCCGTCTT GGAAACCTG GCCTTCTGG AGCCGGGGT CTCACTGGCC GCCCTGGTGA 1860
 TGCTGGTCTT CAAGGCAAAG TTGGCCCTTC TGGAGCCCTT GGTGAAGATG GTCGTCTGG 1920
 25 ACCTCAGGT CCTCAGGGGG CTCGTGGGCA GCCTGGTGTG ATGGGTTTCC CTGGCCCCAA 1980
 AGGTGCAAC GGTGACCTG GCAAAGCTGG TGAGAAGGGA CTGCTGGTG CTCCTGGTCT 2040
 GAGGGGTCTT CTGGCAAAG ATGGTGAGAC AGGTGCTGCA GGACCCCTG GCCCTGCTGG 2100
 ACCTGTCTGT GAACGAGGGC AGCAGGGTGC TCCTGGGCCA TCTGGGTTCC AGGGACTTCC 2160
 TGGCCCTCTT GGTCCCCAG GTGAAGGTGG AAAACCAAGT GACCAGGGTG TTCCCGGTGA 2220
 30 AGCTGGAGCC CTGGCCTCG TGGGTCCAG GGTGAACGA GGTTTCCAG GTGAACGTGG 2280
 CTCTCCCGGT GCCCAGGGCC TCCAGGTGCC CCGTGGCTC CCCGCACTC CTGGCACTGA 2340
 TGGTCCCAA GGTGCATCTG GCCCAGCAGG CCCCCTGGC GCACAGGGCC CTCCAGGTCT 2400
 TCAGGGAATG CTTGGCGAGA GGGGAGCAGC TGGTATCGT GGGCCCCAAG GCGACAGGGG 2460
 TGACGTGGT GAGAAAGGCC CTGAGGGAGC CCTGGAAAG GATGGTGGAC GAGGCTGAC 2520
 35 AGGTCCCATT GGCCCCCTG GCCCAGCTGG TGCTAACGGC GAGAAGGGAG AAGTTGGACC 2580
 TCCTGGTCTT GCAGGAAGTG CTGGTCTCG TGGCGTCCG GGTGAACGTG GAGAGACTGG 2640
 CCCCCCGGA CCAGCGGGAT TTGCTGGGCC TCCTGGTGCT GATGGCCAGC CTGGGGCCAA 2700
 GGGTGAGCAA GGAGAGGCCG GCCAGAAAG CGATGCTGGT GCCCTTGGTC CTCAGGGCCC 2760
 CTCTGGAGCA CTGGGGCTC AGGGTCTAC TGGAGTACT GGTCTTAAAG GAGCCGAGG 2820
 40 TGCCCAAGGC CCCCAGGGG CCACTGGATT CCCTGGAGCT GCTGGCCGCG TTGGACCCC 2880
 AGGTCCAAT GGCAACCTG GACCCCTGG TCCCCTGGT CCTTCTGGA AAGATGGTCC 2940
 CAAAGGTGCT CGAGAGGACA GCGGCCCCC TGGCCGAGCT GGTGAACCCG GCCTCCAAGG 3000
 TCCTGCTGGA CCCCCTGGC AGAAGGGAGA GCCTGGAGAT GACGGTCCCT CTGTGCGCA 3060
 AGGTCCACCA GGTCCCCAGG GTCTGGCTGG TCAGAGAGGC ATCGTGGTC TGCTGGGCA 3120
 45 ACGTGGTGAG AGAGGATTCC TTGGCTTGGC TGGCCATCG GGTGAGCCCG GCAAGCAGGG 3180
 TGCTCTGGA GCATCTGAG ACAGAGGTCC TCCTGGCCCC GTGGTCTCT CTGGCCTGAC 3240
 GGGTCTGCA GGTGAACCCG GACGAGAGGG AAGCCCCGT GCTGATGGCC CCCCTGGCAG 3300
 AGATGGCGCT GTGGAGTCA AGGTGATCG TGGTGAGACT GGTGTGTGG GAGCTCTGG 3360
 AGCCCCGGG CCCCCTGGT CCCCCTGGC CGCTGGTCCA ACTGGCAAGC AAGGAGACAG 3420
 50 AGGAGAAGCT GGTGCAAG GCCCATGGG ACCCTCAGGA CCAGCTGGAG CCGGGGAAT 3480
 CCAGGGTCTT CAAGGCCCA GAGGTGACAA AGGAGAGGCT GGAGAGCCTG GCGAGAGAGG 3540
 CCTGAAGGGA CACCTGGCT TCACTGGTCT GCAGGGTCTG CCCGGCCCTC CTGGTCTTC 3600
 TGGAGACCAA GGTGCTTCTG TCCTGCTGG TCCTTCTGGC CTAAGAGGTC CTCCTGGCCC 3660
 CGTGGTCCC CTGGCAAAG ATGGTGCTAA TGGAATCCCT GGCCCCATTG GGCTCTCTGG 3720
 55 TCCCCGTGGA CGATCAGGGC AAACCCGTCC TGCTGGTCTT CTGGAAATC CTGGGCCCCC 3780
 TGGTCTTCCA GGTCCCCCTG GCCCTGGCAT CGACATGTCC GCCTTTGCTG GCTTAGGCC 3840
 GAGAGAGAAG GGCCCCGACC CCCTGCAGTA CATGCGGGCC GACCAGGCAG CCGGTGGCCT 3900
 GAGACAGCAT GACGCGGAG TGGATGCCAC ACTCAAGTCC CTCAACAACC AGATTGAGAG 3960
 CATCCGAGC CCCGAGGGCT CCCGCAAGAA CCCTGCTCGC ACCTGCAGAG ACCTGAAACT 4020
 60 CTGCCACCTT GAGTGAAGA GTGGAGACTA CTGGATTGAC CCAACCAAG GCTGCACCTT 4080
 GGACGCCATG AAGGTTTCT GCAACATGGA GACTGGCGAG ACTTGCTCT ACCCAATCC 4140
 AGCAAACGTT CCAAGAAGA ACTGGTGGAG CAGCAAGAGC AAGGAGAAGA AACACATCTG 4200
 GTTTGGAGAA ACCATCAATG GTGGCTTCCA TTTCAGCTAT GGAGATGACA ATCTGGTCC 4260
 CAACACTGCC AACGTCAGA TGACCTTCTT ACGCTGTCTG TCCACGGAAG GCTCCAGAA 4320
 65 CATCACTAC CACTGCAAG ACAGCATTGC CTATCTGAGC GAAGCAGCTG GCAACCTCAA 4380
 GAAGGCCCTG CTCATCCAGG GCTCCAATGA CGTGGAGATC CGGGCAGAGG GCAATAGCAG 4440
 GTTACGTAC ACTGCCCTGA AGGATGGCTG CACGAAACAT ACCGGTAAGT GGGGCAAGAC 4500
 TGTTATCGAG TACCGGTAC AGAAGACCTC ACGCTCCCC ATCATGACA TTGCACCAT 4560
 70 GGCATAGGA GGGCCGAGC AGGAATTCGG TGTGGACATA GGGCCGTCT GCTTCTGTA 4620
 AAAACCTGAA CCCAGAAACA ACACAATCCG TTGCAAAACC AAAGGACCCA AGTACTTTC 4680
 AATCTCAGTC ACTCTAGGAC TCTGCACTGA ATGGCTGACC TGACCTGATG TCCATTATC 4740
 CCAACCTCTC ACAGTTCGGA CTCTCTCCC CTCTCTTCT AAGAGACCTG AACTGGGCAG 4800
 ACTGCAAAAT AAAATCTCGG TGTCTATTT ATTTATGTC TTCTGTAA ACCTTCGGGT 4860
 75 CAAGGCAGAG GCAGGAAACT AACTGGTGTG AGTCAAATGC CCCGTGAGT ACTGCCCCA 4920
 GCCCAGGCA GAAGACTCTT CTTCAGGTGC CGGGCGCAGG AACTGTGTGT GTCCTACACA 4980
 ATGGTGCTAT TCTGTGTC AAACCTCTGT ATTTTTTAAA ACATCAATTG ATATTAAAAA 5040
 TGAAGGATT ATTGGAAAGT

Protein Accession #: NP_001835.2

1 11 21 31 41 51

5 MIRLGAPQSL VLLTLLVAAY LRCQGGQDVQE AGSCVQDQQR YNDKDVWKPE PCRICVCDTG 60
 TVLCDDIICE DVKDCLSPEI PFGECCPICP TDLATASGQP GPKGQKGEPPG DIKDIVGPKG 120
 PPGPQGPAGE QGPRGRDRGDK GEKGAPGPRG RDGEPGTGPN PGPPGPPGPP GPPGLGGNFA 180
 AQMAGGFDEK AGGAQLQVVMQ GPMGPMGPRG PPGPAGAPGP QGFQGNPGEPP GEPGVSGPMG 240
 10 PRGPPGPPGK PGDDGEAGKP GKAGERGPPG PQGARGFPGT PGLPGVKGHR GYPGLDGAKE 300
 EAGAPGVKGE SGSPGENGSP GPMGPRGLPG ERGRTGPAGA AGARGNDGQP GPAGPPGPVG 360
 PAGGPFPGA PGAKEAGAPT GARGPEGAQG PRGEPGTPGS PGAGASGNP GTDGIPGAKG 420
 SAGAPGIAGA PGFPGRGPP GPQATGPLG PKGQTGEPGI AGFKGEQGP GEPGAPGPQ 480
 APGPAGEEGK RGARGEPGGV GPFGPPGERG APGNRGFPQG DGLAGPKGAP GERGPSGLAG 540
 PKGANGDPGR PGEPGLPGAR GLTGRPGDAG PQGKVGPSA PGEDGRPGPP GPQARGQPG 600
 15 VMGFPKPGA NGEPGKAGEK GLPGAPGLRG LPGKDGETGA AGPFGPAGPA GERGEQGPAG 660
 PSFGQLGPG PGPPGEGGPK GDQGVPEAG APGLVGRPE RGFPGERGSP GAQGLQGP 720
 LPGTPGTDPG KGASGPAGPP GAQGPGLQG MPGERGAAGI AGPKGDRGDV GEKGPEGAPG 780
 KDGGRGLTGP IGPPGAGAN GEKGEVGP PGASAGARGA PGERGETGPP GPAGFAGPPG 840
 20 ADGQGAKEG QGEAGQKGA GAPGQGPSG APGQGGTGV TGPKGARGAQ GPPGATGFP 900
 AAGRVPGPS NGNPMPGPP GPSGKDGPKG ARGDSGPPGR AGEPLQGP GPPGEKGEPP 960
 DDGSGAEGP PGPPGLAGQR GIVGLPGQR ERGFPGLPG SEEPGKQGP GASGDRGPPG 1020
 PVGPPGLTGP AGEPPREGSP GADGPPGRDG AAGVKGDRGE TGAVGAPGAP GPPGSPGAP 1080
 PTGKQDRGE AGAQGPMGPS GPAGARGIQG PQGPRGDKGE AGEPPGERGLK GHRGFTGLQG 1140
 25 LPGPPGSGD QGASGPAGPS GPPGPPGPV PSKDGANGI PGPPGPPGR GRSGETGPAG 1200
 PPGNPPGPPG PGPPGPGIDM SAFAGLGP RE KGPDPQYMR ADQAAGGLRQ HDAEVDATLK 1260
 SLNNQIESIR SPEGRKNA RTCRDLKLC PEWKSQDYWI DPNQGCITLDA MKVFCNMETG 1320
 ETCVYPNPAN VPKNWSSK SKEKKHIWFG ETINGGFHFS YGDDNLAPNT ANVQMTFLRL 1380
 30 LSTEGSQNIT YHCKNSIAYL DEAGNLKKA LLIQGSNDVE IRAEGNSRFT YTALKDGGCTK 1440
 HTGKWGKTVI EYRSQKTSRL PIIDAPMDI GGPEQEFQVD IGPVCF

SEQ ID NO:125 PFH9 DNA SEQUENCE

Nucleic Acid Accession #: NM_005084

Coding sequence: 162-1487 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51

40 GCTGGTCGGA GGCTCGCAGT GCTGTCGGCG AGAAGCAGTC GGGTTTGGAG CGCTTGGGTC 60
 GCGTTGGTGC GCGGTGGAAC GCGCCAGGG ACCCCAGTTC CCGCGAGCAG CTCGCGCCG 120
 CGCCTGAGAG ACTAAGCTGA AACTGCTGCT CAGCTCCCAA GATGGTGCCA CCCAAATTGC 180
 ATGTGCTTTT CTGCTCTGCG GCTGCTGCTG CTGTGGTTTA TCCTTTTGAC TGGCAATACA 240
 TAAATCCTGT TGCCCATATG AAATCATCAG CATGGGTCAA CAAAATACAA GTACTGATGG 300
 45 CTGCTGCAAG CTTTGGCCAA ACTAAAATCC CCCGGGGA AAA TGGGCTTAT TCCGTTGGTT 360
 CTGACAGACTT AATGTTTGTAT CACACTAATA AGGGCACCTT CTTGCGTTTA TATTATCCAT 420
 CCTAAGATAA TGATCGCCTT GACACCTTT GGATCCCAA TAAAGAATAT TTTTGGGGTC 480
 TTAGCAAAAT TCTTGAACA CACTGGCTTA TGGGCAACAT TTTGAGGTTA CTCTTTGGTT 540
 CAATGACAAC TCTGCAAC TGGAAATCCC CTCTGAGGCC TGGTGAAAAA TATCCACTTG 600
 50 TTGTTTTTTC TCATGGTCTT GGGGCATTCA GGACACTTTA TTCTGCTATT GGCATTGACC 660
 TGGCATCTCA TGGGTTTATA GTTGCTGCTG TAGAACACAG AGATAGATCT GCATCTGCAA 720
 CTCTACTATT CAAGGACCAA TCTGCTGCAG AAATAGGGGA CAAGTCTTGG CTCTACCTTA 780
 GAACCTGAA ACAAGAGGAG GAGACACATA TACGAAATGA CGAGGTACGG CAAAGAGCAA 840
 AAGAATGTTT CCAAGCTCTC AGTCTGATTC TTGACATTGA TCATGGAAG CCAAGTGAAGA 900
 55 ATGCATTAGA TTAAGATTG GATATGGAAC AACTGAAGGA CTCTATTGAT AGGGAAAAA 960
 TAGCAGTAAT TGGACATTCT TTTGGTGGAG CAACGGTTAT TCAGACTCTT AGTGAAGATC 1020
 AGAGATTCAG ATGTGGTATT GCCCTGGATG CATGGATGTT TCCACTGGGT GATGAAGTAT 1080
 ATTCCAGAAT TCCTCAGCCC CTCTTTT TCAACTCTGA ATATTCCAA TATCCTGCTA 1140
 60 ATATCATAAA AATGAAAAA TGCTACTCAC CTGATAAAGA AAGAAAGATG ATTACAATCA 1200
 GGGGTTTCACT CCACCAGAAT TTTGCTGACT TCACTTTTGC AACTGGCAAA ATAATTGGAC 1260
 ACATGCTCAA ATTAAGGGA GACATAGATT CAAATGTAGC TATTGATCTT AGCAACAAAG 1320
 CTTCATTAGC ATTCTACAA AAGCATTTAG GACTTCATAA AGATTTTGAT CAGTGGGACT 1380
 65 GCTTGATTGA AGGAGATGAT GAGAATCTTA TTCCAGGGAC CAACATTAAC ACAACCAATC 1440
 AACACATCAT GTTACAGAAC TCTTCAGGAA TAGAGAAATA CAATTAGGAT TAAATAGGT 1500
 TTTT

SEQ ID NO:126 PFH9 Protein sequence:

Protein Accession #: NP_005075.1

1 11 21 31 41 51

70 MVPPKLHVLF CLCGCLAVVY PFDWQYINPV AHMKSSAWVN KIQVLMMAAS FGQTKIPRGN 60
 GPYSVGCTDL MFDHTNKGTF LRLYYPSQDN DRDLTLWIPN KEYFWGLSKF LGTHWLMGNI 120
 75 LRLFLGSMIT PANWNSFLRP GEKYPLVVS HGLGAFRTLY SAIGIDLASH GFIVA AVEHR 180
 DRASATYTF KDQSAEIGD KSWLYLRTLK QEEETHIRNE QVRORAKECS QALSLILDID 240
 HGKPVKNALD LKFDMEQLKD SIDREKIAVI GHSFGGATVI QTLSEDQRF CRGIALDAWMF 300
 PLGDEVYSRI PQPLFFINSE YFYYPANIK MKKCYSPDKE RKMITIRGSV HQNFADFTFA 360
 TGKIHGMLK LKGDIDSNVA IDLSNKASLA FLQKHLGLHK DFDQWDCLIE GDDENLIPGT 420
 NINTTNQHIM LQNSSGIEKY N

SEQ ID NO:127 PFH8 DNA SEQUENCE

Nucleic Acid Accession #: NM_015900

Coding sequence: 32-1402 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 10 CACGAGCGGC ACGAGGATTT CCAGCTCAGC GATGCCCCCA GGTCCCTGGG AGAGCTGCTT 60
 CTGGGTGGGG GGCCTCATTT TGTGGCTCAG CGTTGGAAGT TCAGGGGATG CACCTCCTAC 120
 CCCACAGCCA AAGTGCCTG ACCTCCAGAG CGCCAACCTT TTTGAAGGCA CCGATCTCAA 180
 AGTCCAGTTT CTCTCTTTG TCCCTTCGAA TCCTAGCTGT GGGCAGCTAG TAGAAGGAAG 240
 CAGTGACCTC CAAAACCTCT GGTTCATGTC CACTCTGGGA ACCAACTAA TTATCCATGG 300
 15 ATTCAGGGTT TTAGGAACAA AGCCTTCCTG GATTGACACA TTTATTAGAA CCCTTCTGCG 360
 TGCAACGAAT GCTAATGTGA TTGCCGTGGA CTGGATTTAT GGGTCTACAG GAGTCTACTT 420
 CTCAGCTGTG AAAAAATGTA TTAAGTTGAG CCTCGAGATC TCCCTTTTCC TCAATAAACT 480
 CTGGTGTCTG GGTGTGTCGG AATCCTCAAT CCACATCATT GGTGTTAGCC TGGGGGCCCA 540
 CGTGGGGGCG ATGGTGGGAG AGCTCTTCGG AGGCCAGCTG GGACAGATCA CAGGCCCTGA 600
 20 CCCCCTGGA CCTGAGTACA CCAGGGGCCAG TGTGGAAGAG CGCTGGATG CTGGAGATGC 660
 CCTCTTCGTG GAAGCCATCC ACACAGACAC CGACAATTTG GGTATTCGGA TTCCCGTTGG 720
 ACATGTGGAC TACTTCGTCA ACGGAGGCCA AGACCAACCT GGCTGCCCCA CCTCTTTTAA 780
 CCGAGTTAT AGTTATCTGA TCTGTGATCA CATGAGGGCT GTGCACCTCT ACATCAGCGC 840
 CCTGGAGAAT TCCTGTCCAC TGATGGCCTT TCCTGTGCC AGCTACAAGG CCTTCCTTGC 900
 25 TGGACGCTGT CTGGATTGCT TTAACCCCTT TCTGCTTTCC TGCCCAAGGA TAGGACTGGT 960
 GGAACAAGGT GGTGTCAAGA TAGAGCCGCT CCCCAAGGAA GTGAAAGTCT ACCTCCTGAC 1020
 TACTTCAGT GCTCCGTACT GCATGCATCA CAGCCTCGTG GAGTTTCACT TGAAGGAACT 1080
 GAGAAACAAG GACACCAACA TCGAGGTTAC CTTCCTTAGC AGTAACATCA CCTCTTCATC 1140
 TAAGATCACC ATACCTAAGC AGCAACGCTA TGGGAAAGGA ATCATAGCCC ATGCCACCCC 1200
 30 ACAATGCCAG ATAAACCAAG TGAAATTCAA GTTTCAGTCT TCCAACCGAG TTGGGAAAAA 1260
 AGACCGGACT ACCATTATTG GGAAGTTCTG CACTGCCCTT TTGCCTGTCA ATGACAGAGA 1320
 AAAGATGGTC TGCTTACCTG AACCAGTGAA CTTACAAGCA AGTGTGACTG TTTCTGTGA 1380
 CCTGAAGATA GCCTGTGTGT AGTTTAACCT GGGCAGGACA CATCTCCCTG CATTTTITTT 1440
 35 TTTTITTTT GAGAGAGAGG TGTGATGAGG GATGTGTGTG TGCAGCTTAT TGTAGACCAT 1500
 TACTACTAAG GAGAAAAGCA AAGCTCTTTC TTATTTTCCT CATAATCAGC TACCCTGGAG 1560
 GGGAGGGAGA ACTCATTTTA CAGAAGTTGG TTTCTTTTGC CGATCTTATG TACATACCCA 1620
 TTTTAGCTTT CCCATGCATA CTAACTGCA CTGTCTTAT CTCCTIGGGC ATTCTACTT 1680
 AGGATTCAAT AGAAACATGT ACAGGGTAAA CAATTTTTTA AAAATAAAAC TTCATGGAGT 1740
 40 AAAAAAAAAA AAAAAAAAAA

SEQ ID NO:128 PFH8 Protein sequence:

Protein Accession #: NP_056984.1

1 11 21 31 41 51
 | | | | |
 45 MPPGPWESCF WVGGLILWLS VGSSGDAPPT PPKCADFQS ANLFEGTDLK VQFLLFVPSN 60
 PSCGQLVEGS SDLQNSGFNA TLGTLKLIHG FRVLGTPKPSW IDTFIRTLR ATNANVIAVD 120
 50 WYGGTGVYF SAVKNVIKLS LEISLFLNKL LVLGVSESSI HIIGVSLGAH VGGMVGQLFG 180
 GQLGQITGLD PAGPEYTRAS VEERLDAGDA LFVEAIHTDT DNLGIRIPVG HVDYFVNGGQ 240
 DQPGCPITFFY AGYSYLCIDH MRAVHLYISA LENSCLMAF PCASYKAFLA GRCLDCFNPF 300
 LLSCPRIGLV EQGGVKIEPL PKEVKVYLLT TSSAPYCMHH SLVEFHLKEL RNKNDTNEVT 360
 55 LFSNITSSS KITPKQQRQY GKGIHAHATP QCQINQVKFK FQSSNRVWKK DRTTHIGKFC 420
 TALLPVNDRE KMVCLPEPVN LQASVTVSCD LKIACV

SEQ ID NO:129 PFH7 DNA SEQUENCE

Nucleic Acid Accession #: NM_014384

Coding sequence: 89-1336 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 65 CGTTGCCGGG TCOCAGGTCC CGCCAGTGCG AGCGCAACGG AGGTGGAAGG CGTTCAGACT 60
 CTTAAGCTGAA CGCGGAGCTG CGGCGGCTAT GCTGTGGAGC GGCTGCCGGC GTTTCGGGGC 120
 GCGCCTCGGC TGCTTGCCTG GCGGTCTCCG GGTCTCTGTC CAGACCGGCC ACCGGAGCTT 180
 GACCTCCTGC ATCGACCTCT CCATGGGACT TAATGAAGAG CAGAAAGAAT TTCAAAAAGT 240
 70 GGCCTTTGAC TTTGCTGCC GAGAGATGGC TCCAAATATG GCAGAGTGGG ACCAGAAGGA 300
 GCTGTTCCTA GTGGATGTGA TGCAGGAGGC AGCCAGCTA GGCTTCGGAG GGGTCTACAT 360
 ACAAACAGAT GTGGCGGGT CTGGGCTGTC ACGTCTTGAT ACCTCTGTCA TTTTGAAGC 420
 CTTGGCTACA GGCTGCACCA GCACCACAGC CTATATAAGC ATCCACAACA TGTGTGCTG 480
 GATGATGAT AGCTTCGGAA ATGAGGAACA GAGGCACAAA TTTGCCAC CGCTCTGTAC 540
 75 CATGGAGAAG TTTGCTTCT ACTGCTCAC TGAACCAGGA AGTGGGAGTG ATGCTGCCTC 600
 TCTTCTGACC TCCGCTAAGA AACAGGGAGA TCATTACATC CTCAATGGCT CCAAGGCCTT 660
 CATCAGTGGT GCTGTGAGT CAGACATCTA TGTGGTCATG TGCCGAACAG GAGGACCAGG 720
 CCCCAAGGGC ATCTGTGATA TAGTTGTGA GAAGGGGACC CCTGGCTCA GCTTTGGCAA 780
 GAAGGAGAAA AAGGTGGGGT GGAAGTCCCA GCCAACACGA GCTGTGATCT TCGAAGACTG 840
 TGCTGTCCCT GTGGCCAACA GAATTGGGAG CGAGGGGACG GGCTTCTCTA TTGCCGTGAG 900

AGGACTGAAC GGAGGGAGGA TCAATATTGC TTCTGCTCC CTGGGGGCTG CCCACGCCTC 960
 TGTCATCCTC ACCCGAGACC ACCTCAATGT CCGGAAGCAG TTTGGAGAGC CTCTGGCCAG 1020
 TAACCAGTAC TTGCAATTCA CACTGGCTGA TATGGCAACA AGGCTGGTGG CCGCGCGGCT 1080
 GATGGTCCGC AATGCAGCAG TGGCTCTGCA GGAGGAGAGG AAGGATGCAG TGGCCTTGTG 1140
 CTCCATGGCC AAGCTCTTTG CTACAGATGA ATGCTTTGCC ATCTGCAACC AGGCCTTGCA 1200
 GATGCACGGG GGCTACGGCT ACCTGAAGGA TTACGCTGTT CAGCAGTACG TGCGGGACTC 1260
 CAGGGTCCAC CAGATTCTAG AAGGTAGCAA TGAAGTGATG AGGATACTGA TCTCTAGAAG 1320
 CCTGCTTCAG GAGTAGAACC CACACTTGTT CTGGCCTGGT GTTCAGTGC GACTGCAGTCA 1380
 GTGTTGAGTG GTGCCATGTG GGCCGCTCTA TTCCAAAGGA ATCATGGATT AGACCCAAAG 1440
 GCTGAGCTCC TCTAGGGCAG GACCTGCACC CTGTGTGTG GCACCAGCAT CGGGTCTTGG 1500
 ACTGGGGCAG AATCCCCAGT GGAACCGGAA GAGCTGGACT GATGAGAAAC ATCAGAAGAA 1560
 CACATACTAC CTGTGTTTCC TAATGCCAGA AGGGTGACCA GTGAAGATT ACCGTCAAAC 1620
 CATGAAAGTC CTTTCTTGGG TCCACTTTAT CTTGATTAGT CTGCATTTTA CTAGTTCACT 1680
 GGATCCCTCC TCTAGGGGCC TGGGGACTTT CACTGATGCT CTTCTGATT CTAGAGCAAA 1740
 GGTGTGGGAA GGGGAAATGG AGGAATGCC TCTGTCTGT GTCGTTCTCT GTGCCACAGC 1800
 TACAGATGCA GAAGGTTTCT CTGGATAGCA CACCTCTGAA TGTAATCAT GATAAAATGG 1860
 ATATTTGAA ACTTACTCCT AAGCTGTGAT GTAGGGTGTA TTCTACTTC TGGACTGCCT 1920
 CAATATCAAG GGTCTGAGCT TTGAATGTT GAATATTCTG TGGGTTTCAT GTTAAGACGC 1980
 CTGTGTCCA GGAAGTGCTAT TCAGTGTTC TGTCTCTGAT AACACTTTG AATATTTT 2040
 TGTGTTTTTG TTCTCTTTC TGAAGCTGTT CCTCTTTTA AATATTTTA ATCATTGA 2100
 TAAATCTAT CCTTCATCCA CCTCTGGTTC TACTATAGT GATTTTATT TAAATGTT 2160
 AATTGTATT GATTAAACAC TTAAGTGGAT TTGGAATAA TAAACTCTC GTCCAATTG 2220
 GCTTTTAAAA AAAAAAAA

SEQ ID NO:130 PFH7 Protein sequence:
 Protein Accession #: NP_055199.1

1 11 21 31 41 51
 | | | | |
 MLWSGCRREF ARLGCLPGGL RVLVQTGHR LTSCIDPSMG LNEEQKEFK VAFDFAAREM 60
 APNMAEWDQK ELFPVDVMRK AAQLGFGGVY IQTDVGGSL SRLDTSVIFE ALATGCTSTT 120
 AYISIHNMCA WMDSFGNEE QRHKFCPLC TMEKFASYCL TEPGSGSDAA SLLTSAKKQG 180
 DHYILNGSKA FISGAGESDI YVVMCRTGGP GPKGISCIIV EKGTPGLSFG KKEKKVGVNS 240
 QPTRAVIDED CAVPVANRIG SEGQGLIAV RGLNGGRINI ASCSLGAHA SVILTRDHLN 300
 VRKQFGEPLA SNQYLQFILA DMATRLVAAR LMVRNAAVAL QEERKDAVAL CSMAKLFATD 360
 ECFaicnQAL QMHGGYGYLK DYAVQYVRD SRVHQILEGS NEVMRILISR SLLQE

SEQ ID NO:131 PFH6 DNA SEQUENCE

Nucleic Acid Accession #: NM_013989
 Coding sequence: 707-1105 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 GCCTGCAGAG AGAGGCATT TGCACCACAG ACAGATAGCA AGAAGGGAAA GACAGAGAGT 60
 GAGAAAAAAG AGGAGTCAGT CGCTCCTGGG GAAGGGAGAG AGTGAGACTG GGAGAAAGAG 120
 AAGCACAGAA AGTGTGTGTA AAACGGAGTA AAGAAAGAAA AAAAAAAAAC TACCCTTAAA 180
 GCACATTTAA AAAAAAAA CTCTGGCAAT TCAAGAAAGA AACAGGCTAC GTTTAAAGAG 240
 CATAGAGACA ATGAAAGGCT AAAGAAAATT TAAAATCTC TGCCACAGTC TCATAGGTGC 300
 TTGGAAATGA AAGTAGAACT GCCTGTCTTT AACGGACTCT GACAGAGGTA ACTGGATTAG 360
 GGACGAGTAC GCCAGCTTTT TTTTCTTTT TTTTCTTTT TTTAATCTC TAAATCCTGA 420
 AAAAAAAGAA AAAAAAAGAA AAAAGGCAGC AGCTCCGAAT TGAATGAATT GATGGGCACA 480
 CTCCAAGTGC TGGGCTGGAG AGACTGGACT TAGTCTTGCC ATTTCTGCTT CTTTGAAAGA 540
 GGAGACAACT TGGGCTTCTT TTTAATTTAG TTTTCTTCC CTTCTCCCC CAACCCCAA 600
 CCTTCCCCCT TACCTCCCCC ACCCCCTTTA TCACCACCCC CCTTTAAAT AAGAGGGTGA 660
 AGGGGAACCA GAGCGACAAA GGGAACTGAC TCAGGAGGCA GAGAAGATGG GCATCCTCAG 720
 CGTAGACTTG CTGATCACAC TGCAAAATCT GCCAGTTTTT TTCTCCAAC TGCCTCTCCT 780
 GGCTCTCTAT GACTCGGTCA TTCTGCTCAA GCACGTGGTG CTGCTGTGA GCCGCTCCAA 840
 GTCCACTCGC GGAGAGTGGC GGCAGATGCT GACCTCAGAG GGAAGTGGCT GCGTCTGGAA 900
 GAGCTTCTC CTGATGCCT ACAACAGGT GAAATTGGGT GAGGATGCC CCAATTCCAG 960
 TGTGTGTCAT GTCTCCAGTA CAGAAGGAGG TGACAACAGT GGCAATGGTA CCCAGGAGAA 1020
 GATAGCTGAG GGAAGCCACT GCCACCTTCT TGACTTTGCC AGCCCTGAGC GCCCACTAGT 1080
 GGTCAACTTT GGCTCAGCCA CTGACCTCC TTTACGAGC CAGCTGCCAG CCTTCCGCAA 1140
 ACTGTGGGAA GAGTCTCTCT CAGTGGCTGA CTCTCTGCTG GTCTACATTG ATGAGGCTCA 1200
 TCCATCAGAT GACTGGGCGA TACCGGGGGA CTCCTCTTTG TCTTTGAGG TGAAGAAGCA 1260
 CCAGAACCAG GAAGATCGAT GTGCAGCAGC CCAGCAGCTT CTGAGCGTT TCTCTTGGC 1320
 GCCCCAGTGC CGAGTTGTGG CTGACCGCAT GGACAATAAC GCCAACATAG CTTACGGGGT 1380
 AGCCTTGAA CGTGTGTGCA TTGTGCAGAG ACAGAAAATT GCTTATCTGG GAGGAAAGGG 1440
 CCCCTTCTCC TACAACCTTC AAGAAGTCCG GCATTGGCTG GAGAAGAATT TCAGCAAGAG 1500
 ATGAAAGAAA ACTAGATTAG CTGGTTAAAG GTATGATTAT AAGAGAGCTT ATTGTTTAA 1560
 AAAGTTATAT AAAGGCAAGG AAATTAAGAA CTGAATCCAT ATTTCAACAG AGCCCTATTG 1620
 GCTTACTGAA AGACAGGAGT TTATCTATCG GAAGAACATG AATCTCTAAC AGCTCCATAC 1680
 TTCTTTCACT ACTCAAAATG CATTGGGCTG AGTAAGTAAC CATATCACCT CTCTCTTAG 1740
 TAAAAAGCCC TATGTGAAAA GATCCCAAGA TGGAGAGGAA GAAACGCTAA TTCAGCATGT 1800
 GTTCATTCTG CATTGAGAAG GAAGTATAC ATCTGATGCA TGCTTTGAGA CCAGAAGAAA 1860
 AGACTTACCT GAATAATTAC TACATTAGGG AAGCTACTGT CTACGTTAAG ATAAAGGGTA 1920

5 TTGCCTTGGC TCTATTTGGC ATGGATGGAG CCCAGTTGGA AAATTCCTCAA ATATTACAAC 1980
 AAGTCCTTGA ACCCAGGCCA TGTGGTTAGA CGTTGGTGT AAGGTTAGAC CTATGTTAG 2040
 AGTCATTCT GATGTTCCAG CTCTAGCCA TGTAGTGCTC TCACTCTCA TACCCAGAA 2100
 ATTATTGGTA TATTGTAGA TACCGAGAAAT GATCCCTCAG TCTGAGAGGT TAGAATGATC 2160
 ATCTGTAATC TGAGGGTTAA TTTCTAGGCA GGTGGAGAGA GTGGTAAAAA AGAAATGAAA 2220
 10 TTGACAAGCT AGGAAAGAGG AGGCAGAAAG ATTGGGAAAA TTCACAGAGT TTCACCCCTA 2280
 AGCTGTAGAG AGTGGGTAC ATTGTGTAGC CACGGAAACA TAGAAACATA CACAAGGCCA 2340
 GAAAAAGAAG AAGGAGCTCA ACTAAAAGTG GCATAGAGAA TACACATATA AAAACAATAT 2400
 ATTTGTCATA TGCTCCTAGA GAGGAGAAAG GGGTGATTGA AAGAAAAAAA AATACCTAAA 2460
 TATTTGTAAT TGTGAGGGGT TCTTTTGGGA AATAATTACT TTGAACCAT GTATGTGGTA 2520
 TGTATATTTT CAGTGGGTTA ATTATACCCC ATGATACCTA TTAAAGGAAA ACCAGTGGGT 2580
 CTGGTGGTGC TGGTCTTTC CTCCCATTC CTACAATTC TATGTGGCCC AAGTCATTCC 2640
 TAATCTTGGT CTCTATAGCA GTGTTCTCTC TGAATGCTGA GCTGAAGAAA TTATACGTAC 2700
 15 ATACACACAT ACATACATAC ATACAAATAT ATGTATATAT ATCTCAGCT GCTGCGGGAG 2760
 GTAGGTACCA TGGCCATTCA GCACAGCCTT GATTTCTCTC CAAAGTAGGT GAGCTATAGT 2820
 GAAGAATAGG TGCAACAAA CAAGCTTACT TCCATTGCAA AATAGAAGAA GAGGAAGTTA 2880
 GAGATAATTC TGATCAATCA TTTTGGAGGC TTTGTTATAA GGCAACCCCC GGTATATCAT 2940
 GGAATTTCCA TTGACATTG AATTGGACT TGGATCTTCC CTGGTCCCA TTAGCTGAGG 3000
 TTTAGTAATC TAAAGTCCCT ATAGTATATG ATTATAATGC TATTTTAAAA AATATATATA 3060
 20 TAAAATATT TTTCTTTT AAAATAGACA CTATAGTTT ACCCATAAGT AATATTTAAA 3120
 GATTATAGCT CCAAAAAGAA TGGACCAACC ACTTTCGTAT CATAATTTCT TTTTGGTAAA 3180
 TATGAGACTA TTATGAAATC ATAGTATATG ATTGTATTTA AAGGTACAAT CAAAGGATCT 3240
 TTTGTCCATT CCATTAATAA CTGAATAAAA AATAAATAAA ATGGATAGAA AAAAATAAAA 3300
 GTTGAAAAATA CATCTTAAA CTAGTTGTCT GAAATGAGAA AAGAGTGAGA ACTAGGTGTG 3360
 25 CAAGAACCAA ACGTATTTTA TTTTATTTT TAAATGGGAG CAACATATCA GTCGTGTCAC 3420
 CAGCTGGTAT ATTGTGTAAT TATTAAAGCT CCATTGGGAC TGATTTTTC TGGCAACATC 3480
 AGCTTTCTAA TGTCTTAAAT TCTATAAAAA CCACCCACAA AGAAACAAAG CAAATTTTCA 3540
 TATCTAATGA GTTCTGGAA AATCATATTG AGAATAATTA TTTGAGATT CTCAGTTGTT 3600
 AACTTCTACA TTCAAGGGCT TATCTCTGCC CCCATTGATT TTAACTCTCA AATGGGTGTG 3660
 30 AGATTACTG TGGAAACCTA AAGCAGTAAA ATAAAAAACC TGGTTGCAGC ACATTCACAC 3720
 TGTGTCTCT AAAATTCCCC TTTTCTCT ATGTACGATA AAGTAACAGT ATGTCAGATA 3780
 AGCCGGTGGG GGGATGAGAT TAGGCTGAGG CAGTGCTAGT CAACTGGGG AAAAGGATGA 3840
 TGGAAAAATC ACCCAGTTGT GCTATATTT TAAAGAAGGA GGTCGTTTAT GTGTGCAGAC 3900
 35 AATTCTCCCT GAGGTAGGCC CAATGGAGAA ATGAAGCAGA GGAAGGAAAC ATAGAAAGAC 3960
 ATGGGCTATC AGGAGGAAG ATGTTCAATA GAACATGCAA GAATTTCTGG AAGAAAGGCT 4020
 GTGGAAGGCG CAATGGAGAA AATGAATGGA CAAAGCTCAG GAATCCCTAC GCTATGTAGA 4080
 ATGTTCTGG TGTATCAGG GTTAAGCCCT GTAATTATGT AACCTATTTA TCGCAACATG 4140
 AATTTTATG ATTTCTGTG ATGATTTCTT TTATGAAATT AACAAGAACT CATTATTTT 4200
 40 AGGTAGAGGA AATCAATGC TTTATCTGAT ATGCTGAGAA ATTATTAGAT TGCCAATACT 4260
 CATGTGCGTT TCATGTGTTT TATAAGGTTT GTTCTTTGA AGAATTGTAG TTCTTAGTCC 4320
 CACAGGAAAA TGTGTATCTA TTTATATATC ATAGTATAAA TCTATGATAT ATTTATATCA 4380
 TATATAAAAG TCTAGTTGTA TTTCTTAGT CCCTAATCAT GTTCTCTCCA TAGGCTGTGT 4440
 TTACATGGAG CTATCGGTTT AGCCTTTTAA GCTTCATTAG CTGTCTATT ATTGAAATAG 4500
 45 TTCCAAGAA ATTTAGATA TTATCATAAC ATCTGGGTCT ACTCAAACAC TTATTGTTT 4560
 AAAGACTTAT GTCTTGGAC TTCAAAAAAC TGACTTTATT TATTGCTTAG TGAAAAACT 4620
 AGTGGGATCA ACAATGATT TCTTGAATGG GCATGAATGG AGATGCCCGC ACAGTAATGT 4680
 AGAAATGTTT CATACATGTA TAAAATGTA ACTGACCTCC TTAGAGGCAG ATTAGTAAT 4740
 GTTCTACTT TGTATAGCTA AGTGACAGTC ACTTAACTTA CATGACTTTC TTTTTCACA 4800
 50 TTGGGTCTCT GGTCTGTGT CTTCACCTCA TTTATAGCAC GTCTCCTTGA TTTTGGTAG 4860
 TATCAACTTC CAGTGATCT GTTCAGTTAA GTTCTTCTCC CGTTAACCAG GAAGTGCTTA 4920
 TTCTCTATC ACAGTGGGAA GAATAGCCTA TTGTCTTCA TTTTGCTGA GTGTATTTA 4980
 CTATTTGGG TCTGAAATAA AAATTATGAA ATATGGTGAG GTCACATGTT GGTGCTGCCT 5040
 TGTGCTATAA AATTCTAGGA GGGCAGTTA GGAGACAGTT ATGTATGGCC TTTCGGGAAA 5100
 55 ATTCAAAGGG TGGGATTACA AGGGTGTTC TCAGGCATGC CCTATGGGC CCTATGTGGA 5160
 AGCAAGAAGA ATTGACTGAT TTACAGGACT TCTCTTTATG TCAATCTTAA GAGGATGGAT 5220
 GAATCTGGAC ATTTGTTCCT CCCAGCTCT GACTGATGGT TTGGAAAAA ACTTAAATTA 5280
 GGATCATATG ACCATTGAAA AAGGAAAAAT GTAGACTCTG ACTTCCGTCC CACTGAAGGA 5340
 60 TTAATGAAAA CCTTACTAG CATTTAGAGC TTTTCAGAAC ATCCCCACTG TCATGTGTCT 5400
 CAGCAGTGA GACTGCAAGT AAGGCTTTA ATTTAGGAG GTTTTTTTTT TTTTTTTTTT 5460
 TTCCCTAAA TGGTATGGCC AAAAGTCAGA GTTAAAAAT ATATAGTTAG ATTCCAACTT 5520
 CCTCTCTAC TCAAAAAATA GAATCCAAAC CCACTCTTCA TATATGCTTC CAGAATGGGG 5580
 CTTAAGTACC AATCTGTCT TTGCAATGGG CACAATCTTG GTCATGTCT GAGGCTCTCT 5640
 65 AAGAAAAGAG AGGATCTAGG ATGGGAGAGC TAGAAAGTTG TAACTGGGA AGAACAAGGC 5700
 CCTGAGGGGT TGGTCTACCA ATCTGGGAAG ATTTGAAAAA AAATCTCTCG CAACTGAAGG 5760
 AAGGCTGAAG GCTGCTGCAA GTCATTGAGT GACTTTAGGA TGAGCAAAAC ATTTGGGCCAC 5820
 TTCTAATGC CCTATGTGTA TAGTACCAGA AGCAAGGTCT CAGACTTAAC AGACCCAGCT 5880
 CTGTCCAAG GTGAGTCTGA ACCAATAGAA AGCAACATG TGCAGATATC CAAACAAGAC 5940
 TGCTCATGCA AGTCGGGGCT GGCTACCCGT CTTAGGCAGC AACAGCAGAG CTCCAGGGAG 6000
 70 TTTATTTCAAT TTTTACTAG ACTTCGAAGA CCCAGCAGAT GTTAAATGAA GTCACTATT 6060
 TGGCTCAAA CTTCCACTTC TCCCTCTCC CTCAAAAAGC CAACAGGTAA ACACATAAAT 6120
 GAAAGAAACC CACAGAAGGG GATGGGAAAT AAAGAAAAAT CTCTCAAGAC TTCTCCAGGC 6180
 CCATGTCACT GGTGAGCGTG GTTTTTATGT GTATTAGGAT TGGGGGATGT GAAGAAATAA 6240
 GTATCCAGTA CTTTATAACC AAAGCAATTA AATGATATTG GGGTAGGGAA TGTGTGGCCAG 6300
 75 TTTTGTTTAG TTTTGCATC ACATTGTCAC CCAGACCTCA CTAGCCCCA AGTAATCGGG 6360
 CGCCCCGAAG AGGAGAGAC AGATGTGCCA GAGTTGACCC AGTGTGCCGA TGATAACTAC 6420
 TGACGAAAGA GTCATCGACC TCAGTTAGTG GTTGGATGTA GTCACATTAG TTTGCTCTC 6480
 CCACTCTTGT TCTCTCTGCC AAGGAGAATA TCGGGGACAT GATGCTAAGA GCCCTGGGTA 6540
 AATGTGGTGA GAATGCACGC GTGCATATGC TACACATATG TGCTTCTCAG TTGCAGAAAA 6600
 TGAAGTCTT TGGGAGATTA TCAGTAGAAA GAGTGTATC ATATTGGTGC TGAGTGCTAT 6660

GTGTGCTTAT ACAATTTGTT CTTGTATTTT AATAAACTTT GAATAAAAGA ATAAAAAAAA 6720
AAAAAAAAAAAA AAAAA

5 SEQ ID NO:132 PFH6 Protein sequence:
Protein Accession #: NP_054644.1

1 11 21 31 41 51
| | | | |
10 MGLSVDLLI TLQILPVFFS NCLFLALYDS VILLKHVVLL LSRSKSTRGE WRRMLTSEGL 60
RCVWKSFLD AYKQVKLGED APNSSVVHVS STEGGDNSGN GTQEKIAEGA TCHLLDFASP 120
ERPLVVNFGS ATXPFTSQL PAFRKLVEEF SSVADFLVY IDEAHPSDGW AIPGDSLSF 180
EVKKHQNQED RCAAQQLLE RFSLPQCRV VADRMNNAN IAYGVAFERV CIVQRQKIAY 240
15 LGGKGPFYSN LQEVHRWLEK NFSKRXXKTR LAG

SEQ ID NO:133 PFH5 DNA SEQUENCE

20 Nucleic Acid Accession #: NM_001141
Coding sequence: 72-2102 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
| | | | |
25 CAGGCGTGTG CCAGGGGGGAG CCCCCTCTG CAGCCCTGTG CGCCGTAGAG AGCTGGACTT 60
AGGCTGGCAG CATGGCCGAG TTCAGGGTCA GGGTGTCCAC CGGAGAAGCC TTCGGGGCTG 120
GCACATGGGA CAAAGTGTCT GTCAGCATCG TGGGGACCCG GGGAGAGAGC CCCCCACTGC 180
CCCTGGACAA TCTCGGCAAG GAGTTCCTG CGGGCGCTGA GGAGGACTTC CAGGTGACGC 240
TCCCGGAGGA CGTAGGCCGA GTGCTGTGCG TGCCGCTGCA CAAGGCCGCC CCAGTGCTGC 300
30 CCCTGCTGGG GCCCTGGCG CCGGATGCCT GGTTCGTGCG CTGGTTCCAG CTGACACCGC 360
CGCGGGGGCG CCACCTCTCT TCCCTGTCT ACCAGTGGCT GGAGGGGGCG GGGACCCTGG 420
TGCTGCAGGA GGTGACAGCC AAGGTGTCTT GGGCAGACCA CCACCCTGTG CTCAGCAAC 480
AGCGCCAGGA GGAGCTTCAG GCCCGGCAGG AGATGTACCA GTGGAAGGCT TACAACCCAG 540
35 GTTGGCTCA CTGCCGGAT GAAAAGACAG TGGAAGACTT GGAGCTCAAT ATCAAATACT 600
CCACAGCCAA GAATGCCAAC TTTTATCTAC AAGCTGGCTC TGCTTTTGCA GAGATGAAAA 660
TCAAGGGGTT GCTGGACCGC AAGGGGCTCT GGAGGAGTCT GAATGAGATG AAAAGGATCT 720
TCAACTTCG GAGGACCCCA GCAGCTGAGC ACGCATTTGA GCAGTGGCAG GAGGATGCCT 780
40 TCTTCGCTC CAGTTCCTG AATGGTCTCA ACCCTGTCTT GATCCGCCG TGTCCTACTACC 840
TCCCAAAGAA CTTCCTCGTC ACTGATGCCA TGGTGGCTCT ATTGTTGGGT CCTGGGACCA 900
GCTTGCAGGC TGAGCTAGAG AAGGGCTCCC TGTTCTTGGT GGATCACGGC ATCCTCTCTG 960
GCATCCAGAC CAATGTCTAT AATGGGAAGC CGCAGTTCTC TGCGGCCCA ATGACCTGTC 1020
TATACCAGAG CCCAGGCTGC GGGCCGCTGC TGCTCTCGC CATCCAGCTC AGCCAGACCC 1080
CCGGCCCAAA CAGCCCCATC TTCCTGCCCA CTGATGACAA GTGGGACTGG TTGCTGGCCA 1140
45 AGACCTGGT GCGCAATGCC GAGTTCCTCT TCCATGAGGC CCTACGCGAC CTGCTGCACT 1200
CACATCTGCT GCCTGAGGTC TTCACCTGG CTACCTGCG TCAGTGCCT CACTGCCACC 1260
CTCTCTCAA GCTGCTGATC CCGCACACCC GATACACCT GCACATCAAC ACACTCGCCC 1320
GGGAGCTGCT TATCGTGCCA GGGCAGGTGG TGGACAGGTC CACAGGCATC GGCATTGAAG 1380
GCTTCTCTGA GTTGATACAG AGGAACATGA AGCAGCTGAA CTATTCTCTC CTGTGTCTGC 1440
50 CTGAGGATAT CGGACCCGA GGAGTTGAAG ACATCCAGG CTACTACTAC CGTGATGATG 1500
GGATGCAGAT TTGGGGTGCA GTGGAACGCT TTGTCTCTGA AATCATCGGT ATCTACTACC 1560
CAAGTGATGA GCTGTGCAA GATGACAGAG AGCTCCAGGC CTGGGTGAGA GAGATCTTCT 1620
CCAAGGGCTT GTCATAACAG GAGAGCTCAG GTATCCCTTC CTCAGTGGAG ACCCGGGAAG 1680
CCCTGGTGCA GTATGTACC ATGGTGATAT TCACCTGCTC AGCCAAGCAT GCGGCTGTCA 1740
55 GTGCAGGGCA GTTGTACTCC TGTGCTTGA TGCCCAACCT GCCACCCAGC ATGCAGCTGC 1800
CACCACCCAC CTCCAAAGGC CTGGCAACAT GCGAGGGCTT CATAGCCACC CTCCCACCTG 1860
TCAATGCCAC ATGTGATGTC ATCCTTGCTC TCTGTTGCT GAGCAAGGAG CCTGGAGACC 1920
AAAGGCCCTT GGGCACCTAT CCGGATGAGC ACTTCACAGA GGAGGCCCT CGGCGGAGCA 1980
TCGCCACCTT CCAGAGCCGC CTGGCCAGA TCTGAGGGG CATCCAGGAG CGGAACCCGG 2040
60 GCCTGGTGCT GCCCTACACC TACCTAGACC CTCCCCTCAT CGAGAACAGC GTCTCCATCT 2100
AATCCCAGG GGAACACAGG CCCAGATGAC ATCCCTTTGA CCACATCGCT CTAGGATAAC 2160
TGGCAGCCAG AGAAAAGGAC TCCTCAGAAA AAACAGGCCC CCATGTGCCT CTCCTGGGAC 2220
AACCAGACTC TGTAATCAC CCCCACCACC ATACACACAC AAAAAACAG AAACAAAATC 2280
AAAACAGAGA AACAGAAAA TCTACCAAGA ACAGAGTCTC AGGACAGAAC CACTGAGTCT 2340
65 TTTGGAGGCT CCAAGCTCA AAGTGCCCGC AGAGCCACCC TTGAGGGTTT TGCTAGTTGG 2400
TTTGTGTTTG CGTTTACAGC CGTGGGGGGA AGCACATAAT CCCGCCAG GGCCCACTAG 2460
CATCCACTGA TTGGACCTTA TGGTACCCA ACTCAAGGAC AGCCACCAAG AAGTGGCTGC 2520
CAAAGAGACT GGGCGCAGTG GCTCATGCC ATAATCCCAG CACTTGGGA GATGGAGGCG 2580
70 GGAATAATCAT TTGAGGTGAG AAGTTCAAGG CCAGCTGGA CGACATAGCG AGACTCCACC 2640
TCTACCAAAA AATAAAAAAT AAAAAACAAA AAAAAAAAAA AAAAA

75 SEQ ID NO:134 PFH5 Protein sequence:
Protein Accession #: NP_001132.1

1 11 21 31 41 51
| | | | |
MAEFRVRVST GEAFGAGTWD KVSVSIVGTR GESPLPLDN LGKEFTAGAE EDFQVTLPED 60
VGRVLLLRVH KAPFVPLLG PLAPDAWFCR WFQLTPPRGG HLLFPCYQWL EGAGTLVQE 120
GTAKVSWADH HPVLQQRRQE ELQARQEMYQ WKAYNPGWPH CLDEKTVEDL ELNIKYSTAK 180

NANFYLQAGS AFAEMKIKGL LDRKGLWRSL NEMKRIFNFR RTPAAEHAFH HWQEDAFFAS 240
 QFLNGLNPVL IRRCHYLPHN FPVTDAMVAS LLGPGTSLQA ELEKGSFLV DHGILSGIQT 300
 NVINGKPKQS AAPMTLLYQS PGCGLPLLA IQLSQTGPN SPIFLPTDDK WDWLLAKTWV 360
 RNAEFSFHEA LTHLLHSHLL PEVFTLATLR QLPCHPLFK LLIPHTRYTL HINTLARELL 420
 IVPQVVDRS TGIQIEGFSE LIQRNMKQLN YSLLCLPEDI RTRGVEDIFG YYYRDDGMQI 480
 WGAVERFVSE IIGIYVPSDE SVQDDRELQA WVREIFSKGF LNQESSGIPS SLETREALVQ 540
 YVTMVIFTCS AKHAAVSAGQ FDSCAWMPNL PPSMQLPPT SKGLATCEGF IATLPPVNAT 600
 CDVILALWLL SKEPGDQRPL GTPYDEHFE EAPRSIATF QSRLAQISRG IQERNRGLVL 660
 PYTYLDPLI ENSVSI

SEQ ID NO:135 PFH4 DNA SEQUENCE

Nucleic Acid Accession #: NM_002742

Coding sequence: 236-2974 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 GAATTCCTTC TCTCCTCCTC CTGCCCCCTT CTCTCGCCCT CCTCCTCCTC CTCGCCCTCC 60
 CCTCCCGATC CTCATCCCCT TGCCCTCCCC CAGCCCAGGG ACTTTTCCGG AAAAGTTTITA 120
 TTTTCCGTCT GGGCTCTCGG AGAAAGAAGC TCCTGGCTCA GCGGCTGCAA AACTTTCCTG 180
 CTGCCGCGCC GCCAGCCCCC GCCCTCCGCT GCCCGGCCCT GCGCCCCGCC GAGCGATGAG 240
 CGCCCCCTCG GTCTGCGGC CGCCAGTCC GCTGTGCCC GTGGCGGCGG CAGTGCCTG 300
 AGCGGCGGCC GCATCGGTCC CAGGGTCCGG GCCCGGGGCC GCGCCGTCTT TGGCTCCTGT 360
 CGCGGCCCCG GTCGGGGCA TCTCGTTCCA TCTGCAGATC GGCCTGAGCC GTGAGCCGGT 420
 GCTGTGCTG CAGGACTCGT CCGGGGACTA CAGCCTGCGC CACGTCCGCG AGATGGCTTG 480
 CTCCATTGTC GACCAGAAGT TCCTGAAATG TGGTTTCTAC GGAATGTATG ATAAGATCCT 540
 GCTTTTTCGC CATGACCTTA CCTCTGAAAA CATCCTTCAG CTGGTGAAAG CGGCCAGTGA 600
 TATCCAGGAA GGCATCTTA TTGAAGTGGT CTGTGACGT TCCGCCACCT TTGAAGACTT 660
 TCAGATTGCT CCCCAGCTC TCTTTGTTCA TTCATACAGA GCTCCAGCTT TCTGTGATCA 720
 CTGTGGAGAA ATGCTGTGGG GCTGTGTACG TCAAGGTCTT AAATGTGAAG GGTGTGGTCT 780
 GAATTACCAT AAGAGATGTG CATTTAAAT ACCTAACAAAT TGCAGCGGTG TGAGGCGGAG 840
 AAGGCTCTCA AACGTTTCCC TCACTGGGGT CAGCACCATC CGCACATCAT CTGCTGAACT 900
 CTCTACAAGT GCCCTGATG AGCCCTTCT GCAAAAATCA CCATCAGAGT CGTTTATTGG 960
 TCGAGAGAAG AGGTCAAAT CTCAATCATA CATTGGACGA CCAATTCACC TTGACAAGAT 1020
 TTTGATGTCT AAAGTTAAAG TGCCGCACAC ATTTGTATC CACTCTTACA CCCGGCCAC 1080
 AGTGTGCCAG TACTGCAAGA AGCTTCTGAA GGGGCTTTTC AGGCAGGGCT TGCAGTGCAA 1140
 AGATTGCAGA TTAACTGGC ATAAACGTTG TGCACCGAAA GTACCAAACA ACTGCCTTGG 1200
 GAAAGTGACC ATTAATGGAG ATTTGCTTAG CCTGGGGGCA GAGTCTGATG TGGTCATGGA 1260
 AGAAGGGAGT GATGACAATG ATAGTGAAG GAACAGTGGG CTCATGGATG ATATGGAAGA 1320
 AGCAATGGTC CAAGATGCAG AGATGGCAAT GGCAGAGTGC CAGAACGACA GTGGCGAGAT 1380
 GCAAGATCCA GACCCAGACC ACGAGGACGC CAACAGAACC ATCAGTCCAT CAACAAGCAA 1440
 CAATATCCCA CTCATGAGGG TAGTGAGTC TGTCAAACAC ACGAAGAGGA AAAGCAGCAC 1500
 AGTCATGAAA GAAGGATGGA TGGTCCACTA CACCAGTAAG GACACGCTGC GGAACCGGCA 1560
 CTATTGGAGA TTGGATAGCA AATGTATTAC CCTCTTTCAG AATGACACAG GAAGCAGGTA 1620
 CTACAAGGAA ATTCCTTTAT CTGAAATTTT GTCTCTGGAA CCAGTAAAAA CTTCAGCTTT 1680
 AATTCCTAAT GGGGCCAAT CTATGTGTTT CGAAATCACT ACGGCAAAATG TAGTGTATTA 1740
 TGTGGGAGAA AATGTGGTCA ATCTTCCAG CCCATACCA AATAACAGTG TTCTCACCAG 1800
 TGGCGTTGGT GAGATGTGG CAGGATGTG GGAGATAGCC ATCCAGCATG CCCTTATGCC 1860
 CGTCATTCCC AAGGGCTCCT CCGTGGGTAC AGGAACCAAC TTGCACAGAG ATATCTCTGT 1920
 GAGTATTTCG GTATCAAATG GCCAGATTCA AGAAAATGTG GACATCAGCA CAGTATATCA 1980
 GATTTTTCCT GATGAAGTAC TGGTCTTGG ACAGTTTGGG ATTGTTTATG GAGGAAAAACA 2040
 TCGTAAAAACA GGAAGAGATG TAGCTATTAA AATCATTGAC AAATTACGAT TTCCAACAAA 2100
 ACAAGAAAGC CAGCTTCGTA ATGAGGTGAC AATTCTACAG AACCTTCATC ACCCTGGTGT 2160
 TGTAAATTTG GAGTGTATGT TTGAGACGCC TGAAAGAGTG TTTGTTGTTA TGGAAAAAAT 2220
 CCATGGAGAC ATGCTGGAAA TGATCTGTGC AAGTGAAAAG GGCAGGTGTC CAGAGCACAT 2280
 AACGAAGTTT TTAATTACTC AGATACTCGT GGCTTTGCGG CACCTTCATT TTAATAATAT 2340
 CGTTCACTGT GACCTCAAAC CAGAAAATGT GTTGCTAGCC TCAGCTGATC CTTTTCCTCA 2400
 GGTGAAACTT TGTGATTTTG GTTTTGCCCG GATCATTGGA GAGAAGTCTT TCCGGAGGTC 2460
 AGTGGTGGGT ACCCCCGCTT ACCTGGCTCC TGAGGTCTTA AGGAACAAGG GCTACAATCG 2520
 CTCTCTAGAC ATGTGTGCTG TTGGGGTCAT CATCTATGTA AGCCTAAGCG GCACATTCCC 2580
 ATTTAATGAA GATGAAGACA TACACGACCA AATTCAGAAT GCAGCTTTC TGTATCCACC 2640
 AAATCCCTGG AAGGAAATAT CTCATGAAGC CATIGATCTT ATCAACAATT TGCTGCAAGT 2700
 AAAAATGAGA AAGCGCTACA GTGTGGATAA GACCTTGAGC CACCTTGGC TACAGGACTA 2760
 TCAGACCTGG TTAGATTGCG GAGAGCTGGA ATGCAAAATC GGGGAGCGCT ACATCACCCA 2820
 TGAAAGTGAT GACCTGAGGT GGGAGAAGTA TGCAGGCGAG CAGCGGCTGC AGTACCCAC 2880
 ACACCTGATC AATCCAAGTG CTAGCCACAG TGACACTCCT GAGACTGAAG AAACAGAAAT 2940
 GAAAGCCCTC GGTGAGCGTG TCAGCATCCT CTGAGTTCCA TCTCTATAA TCTGTCAAAA 3000
 CACTGTGGAA CTAAATAAATA CATACGGTCA GGTTTAACAT TTGCCTTGCA GAACTGCCAT 3060
 TATTTTCTGT CAGATGAGAA CAAAGCTGTT AAAGTGTAG CACTGTTGAT GTATCTGAGT 3120
 TGCCAAAGCA AATCAACAGA AGCATTGTGA TTTGTGTGA CCAACTGTGT TGTATTAAACA 3180
 AAAGTCCCTT GAAACACGAA ACTTGTTATT GTGAATGATT CATGTTATAT TTAATGCATT 3240
 AAACCTGTCT CACTGTGCCC TTGCAAAATC AGTGTTTTTC TTAAGTGGAG TTCATTTTGG 3300
 TAAGAGACAG AATGTATCTG TGAAGTAGTT CTGTTTGGTG TGTCCCATTT GTGTTGTCAT 3360
 TGTAACAAAA CTCTTGAGA GTCGATTATT TCCAGTGTTC TATGAACAA TCCAAAACCC 3420
 ATGTGGGAAA AAAATGAATG AGGAGGGTAG GGAATAAAAT CTAAGACAC AAATGCATGA 3480
 ACAAGTTTAA ATGTATAGTT TTGAATCCTT TGCCTGCCTG GTGTGCCTCA GTATATTAA 3540
 ACTCAAGACA ATGCACCTAG CTGTGCAAGA CTAAGTGTCT TTAAGCCTAA ATGCCTTAGA 3600
 AATGAAACTT GCCATATATA CAGATACAT TCCCTCTTT CTATAATAC TCTGTTGATC 3660

TATGAAAAAT CAGCTGCTCA GCAACCTTTC ACCTTTGTGT ATTTTTCAT AATAAAAAAT 3720
ATTCCTGTCA AAAAAAAAAA AA

5

SEQ ID NO:136 PFH4 Protein sequence:
Protein Accession #: NP_002733.1

10 1 11 21 31 41 51
MSAPPVLRPP SPILLVAAAA AAAAAALVPG SGPAPAPFLA PVAAPVGGIS FHLQIGLSRE 60
PVLLQDSSG DYSLAHVREM ACSIVDQKFP ECGFYGMYDK ILLFRHDPIS ENILQLVKAA 120
SDIQEGDLIE VVLSRSATFE DFQIRPHALF VHSYRAPAFC DHCGEMLWGL VRQGLKCEGC 180
GLNYHKRCAF KIPNNCSGVR RRRLSNVSLT GVSRTTSSA ELSTAPDEP LLQKSPSESF 240
15 IGREKRSNSQ SYIGRPIHLD KILMSKVVP HTFVIHSYTR PTVCQYCKKL LKGLFRQGLQ 300
CKDCRFNCHK RCAPKVPNNC LGEVTINGDL LSPGAESDVV MEEGSDDDNS ERNSGLMDDM 360
EAMVQDAEM AMAECQNDG EMQDPDPDHE DANRTISPST SNNPLMRVV QSVKHTKRKS 420
STVMKEGWMV HYTSKDTLRK RHYWRDLKSK ITLFQNDTGS RYYKEIPLSE ILSLEPVKTS 480
20 ALIPNGANPH CFEITTANVV YYVGENVVNP SSPSPNNSVL TSGVGADVAR MWEIAIQHAL 540
MPVIPKGSVV GTGTNLHRDI SVSISVSNQC IQENVDISTV YQIFPDEVLG SGQFGIVYGG 600
KHKRTGRDVA IKIIDKLRFK TKQESQLRNE VAILQNLHHP GVVNLECMFE TPERVFVME 660
KLHGDMLEMI LSEKGRLEP HITKFLITQI LVALRHLHFK NIVHCDLKPE NVLLASADPF 720
25 PQVKLCDFGF ARIEGESFR RSVVGTAYL APEVLRNKG YNRS LDMWSVG VIIYVSLSGT 780
FPNEDEDIH DQIQNAAFMY PPNPWKEISH EADILNNLL QVKMRKRYSV DKTLSHPWLQ 840
DYQTWDLRE LECKIGERYI THESDDLWE KYAGEQRLQY PTHLINPSAS HSDTPETEET 900
EMKALGERVS IL

30

SEQ ID NO:137 PFH3 DNA SEQUENCE

Nucleic Acid Accession #: X95425
Coding sequence: 712-3825 (underlined sequences correspond to start and stop codons)

35 1 11 21 31 41 51
AATGGTCAGT CAATACATTA TAACATAATA CACCAAATGC TAGAATAGAA GGGGAGGGGG 60
GCACACATAA TGACTCACTG CTGGAAGAAG GGTGCATCAG TGAATTAAAA AATGTCCCTC 120
CCCTCTTCAG CACTCAGCGC GCAGCTATT CTTCTGCCA GTCTCTTGA ACTCTGGATC 180
40 TTGCTTTTG CTCGCTGCTC TCCTGTTTTT CATTCTCCAC ATTTCTCAA TCCTCTTTCT 240
TTATCCTTAG CCACCTGCT TTTTCTCTCC TTTTAAAAA AATCGGAGAT TTCGTCTTAA 300
AATGATTGTG CTTCTTACC TTCGTCCATT TCAACACTGA AGGCTGCAAA GAACCTCACC 360
TTTCCCCTAG TGGTATTTAA AAATTCTCAA TCCGTAAAAA GTCTTTTGA AAGGCAAAGG 420
AACAGGACCC AGACCTCTC GACACCTTG ATCCGAGTCA GATCTGCACT AGCAACCAGA 480
45 ACTAATATTT CATTAAACC ACCAAAAGGG GGAGGCGAGA GGAGCCAGAA GCAAACCTCA 540
TCTGTCTCAG ACGGATCCGT GGTTCCTACA TTTGGAGGAG CCGCGTGTA GAAGGCGTAG 600
GACCCCAAGG GGGGACAAGG AGGACTCCCG AGTCTCCCT CTCCGCTCTC CGAGACCGAA 660
GAGGTGGACT GAGCCGCTCG GGACAGCGGC ACCGGAGGAG GCTCGGAGAA GATGCGGGGC 720
TCGGGGCCCC GGGGTGCGGG ACACCGCGGG CCCCAGCG GCGGCGGCGA CACCCCATC 780
50 ACCCAGCGT CCGTGCGCG CTGCTACTCT GCACCTCGAC GGGCTCCCT CTGGACGTGC 840
CTTCTCTGT GCGCGCACT CCGGACCTC CTGGCAGCC CCAGCAACGA AGTGAATTTA 900
TTGGATTCA GCACGTGAT GGGGGACCTG GGATGGATTG CTTTCCAAA AAATGGGTGG 960
GAAGAGATTG GTGAAGTGGA TGAATAATTAT GCCCTATCC ACACATACCA AGTATGCAAA 1020
GTGATGGAAC AGAATCAGAA TAACTGGCTT TTGACCAATT GGATCTCCAA TGAAGGTGCT 1080
TCCAGAATCT TCATAGAACT CAAATTTACC CTGCGGGACT GCAACAGCCT TCCTGGAGGA 1140
55 CTGGGACCT GTAAGGAAAC CTTAATATG TATTACTTTG AGTCAGATGA TCAGAAATGGG 1200
AGAAACATCA AGGAAACCA ATACATCAAA ATTGATACCA TTGCTGCCGA TGAAGCTTT 1260
ACAGAACTTG ATCTTGGTGA CCGTGTATG AAACCTGAATA CAGAGGTCAG AGATGTAGGA 1320
CCTCTAAGCA AAAAGGGATT TTATCTTGCT TTTCAAGATG TTGGTGCTTG CATTGCTCTG 1380
60 GTTCTGTGC GTGTATACTA TAAAAAATGC CCTCTGTGG TACGACACTT GGCTGTCTTC 1440
CCTGACACCA TCACTGGAGC TGATTCTTCC CAATTGCTCG AAGTGTGAGG CTCCTGTGTC 1500
AACCATTCTG TGACCGATGA ACCTCCCAAA ATGCACTGCA GCGCCGAAGG GGAGTGGCTG 1560
GTGCCATCG GGAAATGCAT TGCAGGCA GGATATGAAG AGAAAAATGG CACCTGTCAA 1620
GTGTGCAGAC CTGGGTCTCT GAAAGCCTCA CCTACATCC AGAGCTGCGG CAAATGTCCA 1680
65 CCTCACAGTT ATACCATGA GGAAGCTTCA ACCTCTGTG TCTGTGAAAA GGATTATTTT 1740
AGGAGAGAGT CTGATCCACC CACAATGGCA TGCACAAGAC CCCCCTCTGC TCCTCGGAAT 1800
GCCATCTCAA ATGTTAATGA AACTAGTGTC TTTCTGGAAT GGATTCCGCC TGCTGACACT 1860
GGTGAAGGA AAGACGTGTC ATATTATATT GCATGCAAGA AGTGCAACTC CCATGCAGGT 1920
70 GTGTGTGAGG AGTGTGGCGG TCATGTGTCAG TACCTTCCCC GGCAAGCGG CCGTGAATAA 1980
ACCTCTGTCA TGATGTGGA TCTACTCGCT CACACAAACT ATACCTTTGA GATTGAGGCA 2040
GTGAATGGAG TGTCCGACT GAGCCAGGA GCCCGGCAGT ATGTGTCTGT AAATGTAACC 2100
ACAAATCAAG CAGCTCCATC TCCAGTCACC AATGTGAAAA AAGGGAAAAA TGCAAAAAAC 2160
AGCATCTCTT TGTCTGGCA AGAACCAGAT CGTCCCAATG GAATCATCTT AGAGTATGAA 2220
ATCAAGCATT TTGAAAAAGGA CCAAGAGACC AGCTACACGA TTATCAAATC TAAAGAGACA 2280
ACTATTACTG CAGAGGGCTT GAAACCAGCT TCAGTTTATG TCTTCCAAAT TCGAGCACGT 2340
75 ACAGCAGCAG GCTATGGTGT CTTCACTCGA AGATTGAGT TTGAAACCAC CCCAGTGTTC 2400
GCAGCATCCA CGATCAAAAG CCAGATTCTT GTAATTGCTG TGTCTGTGAC AGTAGGAGTC 2460
ATTTGTGTGG CAGTGGTAT CGGCGTCTC CTCAGTGGA GTTGCTGCGA ATGTGGCTGT 2520
GGGAGGGCTT CTTCCCTGTG CGCTGTGCC CATCCAATCC TAATATGGCG GTGTGGCTAC 2580
AGCAAGCAA AACAAGATCC AGAAGAGGAA AAGATGCATT TTCATAATGG GCACATTAAA 2640

CTGCCAGGAG TAAGAACTTA CATTGATCCA CATACCTATG AGGATCCCAA TCAAGCTGTC 2700
 CACGAATTTG CCAAGGAGAT AGAAGCATCA TGTATCACCA TTGAGAGAGT TATTGGAGCA 2760
 GGTGAATTTG GTGAAGTTTG TAGTGGACGT TTGAACTAC CAGGAAAAAG AGAATTACCT 2820
 GTGGCTATCA AAACCCCTAA AGTAGGCTAT ACTGAAAAGC AACGCAGAGA TTTCTAGGT 2880
 GAAGCAAGTA TCATGGGACA GTTTGATCAT CCTAACATCA TCCATTTAGA AGGTGTGGTG 2940
 ACCAAAAGTA AACCAATGAT GATCGTGACA GAGTATATGG AGAATGGCTC TTTAGATACA 3000
 TTTTGAAGA AAAACGATGG GCAGTTCAC GTGATTACG TTGTTGGCAT GCTGAGAGGT 3060
 ATCTCTGCAG GAATGAAGTA CCTTCTGAC ATGGGGCTATG TGCATAGAGA TCTTGCTGCC 3120
 AGAAACATCT TAATCAACAG TAACCTTGTG TGCAAAAGTG CTGACTTTGG ACTTTCCCGG 3180
 GTACTGGAAG ATGATCCCGA GGCAGCCTAC ACCCAAAGGG GAGGAAAAAT TCCAATCAGA 3240
 TGGACTGCC CAGAAGCAAT AGCTTTCCGA AAGTTTACTT CTGCCAGTGA TGTCTGGAGT 3300
 TATGGAATAG TAATGTGGGA AGTTGTGTCT TATGGAGAGA GACCTACTG GGAGATGACC 3360
 AATCAAGATG TGATTAAGC GGTAGAGGAA GGCTATCGTC TGCCAAGCCC CATGGATTGT 3420
 CCTGCTGCTC TCTATCAGTT AATGCTGGAT TGCTGGCAGA AAGAGCGAAA TAGCAGGCC 3480
 AAGTTTGATG AAATAGTCAA CATGTTGGAC AAGCTGATAC GTAACCCAAG TAGTCTGAAG 3540
 ACGCTGGTTA ATGCATCTG CAGAGTATCT AATTTATGG CAGAACATAG CCCACTAGGA 3600
 TCTGGGGCCT AGCATCAGT AGGTGAATGG CTAGAGGCAA TCAAGATGGG CCGGTATACA 3660
 GAGATTTTCA TGGAAAATGG ATACAGTTCA ATGGACGCTG TGGCTCAGGT GACCTTGGAG 3720
 CAATTGAGAC GGCTTGGAGT GACTCTTGTG GGTACCAGA AGAAGATCAT GAACAGCCTT 3780
 GAGAAAATGA AGGTGAGCT GGTAAACGGA ATGGTGCCAT TGTAACCTCA TGTAAATGTC 3840
 GCTTCTTCAA GTGAATGATT CTGCACTTTG TAAACAGCAC TGAGATTAT TTTAACAAAA 3900
 AAA

SEQ ID NO:138 PFH3 Protein sequence:
 Protein Accession #: CAA64700.1

1 11 21 31 41 51
 MRGSGPRGAG HRRPPSGGGD TPITPASLAG CYSAPRRAPL WTCLLLCAAL RTLLASPSNE 60
 VNLLDSRTVM GDLGWIAFPK NGWEEIGEVD ENYAPIHTYQ VCKVMEQNQN NWLLTSWISN 120
 EGASRIFIEL KFTLRDCNSL PGGLGTCKET FNMYYFESDD QNGRNIKENQ YIKIDTIAAD 180
 ESFTELDLDG RVMKLNTEVR DVGPLSKKGF YLAFQDVGAC IALVSVRVYY KKCPVVRHL 240
 AVFPDITIGA DSSQLLEVSG SCVNHSVTDE PPKMHCSAEG EWLVPIGKCM CKAGYEEKNG 300
 TCQVCRPGFF KASPHQSCG KCPHSYTHE EASTSCVCEK DYFRRESDDP TMACTRPPSA 360
 PRNAISNVNE TSVFLEWIPP ADTGGRKDVY YYIACKKCNH HAGVCEECGG HVRVLPQSG 420
 LKNTSVMMDV LLAHTNYTFE IEAVNGVSDL SPGARQYVSV NVTTNQAPS PVTNVKKGKI 480
 AKNSISLSWQ EPDRPNHIL EYEIKHFED QETSYTIKS KETTITAEGL KPASVYVFQI 540
 RARTAAGYGV FSRRFFETT PVFAASSDQS QIPVIAVSVT VGVILLAVVI GVLLSGSCCE 600
 CGCGRASSLC AVAHPILWR CGYSKAKQDP EEEKMHFHNG HIKLPVGRVY IDPHTYEDPN 660
 QAVHEFAKEI EASCITIERV IGAGEFGEVC SGRLKLPGRK ELPVAIKTLK VGYTEKQRRD 720
 FLGEASIMGQ FDHNPNIHLE GVVTKSPVM IVTEYMENGSL DITFLKKNDD QFTVIQLVGM 780
 LRGISAGMKY LSDMGYVHRD LAARNILINS NLVCKVSDFG LSRVLEDDPE AAYTTRGGKI 840
 PIRWTAPEAI AFRKFTSASD VWSYGIVMWE VVSYGERPYW EMTNQDVKA VEEGYRLPSP 900
 MDCPAALYQL MLDWCQKERN SRPKFDEIVN MLDKLIRNPS SLKTLVNASC RVSNNLAEHS 960
 PLGSGAYRSV GEWLEAIKMG RYTEIFMENG YSSMDAVAQV TLEDLRLRGV TLVGHQKKIM 1020
 NSLQEMKVQL VNGMVP

SEQ ID NO:139 PFH2 DNA SEQUENCE

Nucleic Acid Accession #: NM_016029
 Coding sequence: 78-1097 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 CTGCGATCCC GCAGGGCAGC GACGCGACTC TGGTGCGGGC CGTCTTCTTC CCCCGAGCT 60
 GGGCGTGCGC GGCCGCAATG AACTGGGAGC TGCTGCTGTG GCTGCTGGTG CTGTGCGCGC 120
 TGCTCTGCTC TTTGGTGACG CTGCTGCGCT TCCTGAGGGC TGACGGCGAC CTGACGCTAC 180
 TATGGGCCGA GTGGCAGGGA CGACGCCAG AATGGGAGCT GACTGATATG GTGGTGTGGG 240
 TGACTGGAGC CTCGAGTGGG ATTGGTGAGG AGCTGGCTTA CCAGTTGTCT AAAGTAGGAG 300
 TTTCTTTGT GCTGTCAGCC AGAAGAGTGC ATGAGCTGGA AAGGGTGAAA AGAAGATGCC 360
 TAGAGAATGG CAATTAAAA GAAAAAGATA TACTTGTITT GCCCTTGAC CTGACCGACA 420
 CTGGTTCCCA TGAAGCGGCT ACCAAAGCTG TTCTCCAGGA GTTTGGTAGA ATCGACATTC 480
 TGGTCAACAA TGGTGAATG TCCAGCGTT CTCTGTGCAT GGATACCAGC TTGGATGTCT 540
 ACAGAAAGCT AATAGAGCTT AACTACTTAG GGACGGTGTC CTGACAAAAA TGTGTTCTGC 600
 CTCACATGAT CGAGAGGAAG CAAGGAAAGA TTGTTACTGT GAATAGCATC CTGGGTATCA 660
 TATCTGTACC TCTTTCATT GGATACTGTG CTAGCAAGCA TGCTCTCCGG GGTTTTTTTA 720
 ATGGCCTTCG AACAGAACTT GCCACATACC CAGGTATAAT AGTTTCTAAC ATTTGCCAG 780
 GACCTGTGCA CCAATAATG GTGGAGAAAT CCCTAGCTGG AGAAGTCACA AAGACTATAG 840
 GCAATAATGG AGACCACTCC CACAAGATGA CAACCACTCG TTGTGTGCGG CTGATGTAA 900
 TCAGCATGGC CAATGATTG AAAGAAGTTT GGATCTCAGA ACAACCTTTC TTGTTAGTAA 960
 CATATTGTG CCAATACATG CCAACCTGGG CCTGGTGGAT AACCAACAAG ATGGGGAAGA 1020
 AAAGGATTGA GAACTTTAA AGTGGTGTGG ATGCAGACTC TTCTTATTTT AAAATCTTTA 1080
 AGACAAAAAA TGACTGAAAA GAGCACCTGT ACTTTTCAAG CCACTGGAGG GAGAAATGGA 1140
 AAACATGAAA ACAGCAATCT TCTTATGCTT CTGAATAATC AAAGACTAAT TTGTGATTTT 1200

ACTTTTAAAT AGATATGACT TTGCTTCCAA CATGGAATGA AATAAAAAAT AAATAATAAA 1260
AGATTGCCAT GAATCTTGCA AA

5

SEQ ID NO:140 PFH2 Protein sequence:
Protein Accession #: NP_057113.1

10

1 11 21 31 41 51
| | | | |
MNWELLWL VLCALLLV QLLRFLRADG DLTLWAEWQ GRRPEWELTD MVVWVTGASS 60
GIGEELAYQL SKLGVSLVLS ARRVELERV KRRCLNENL KEKDILVPL DLTDTGSHEA 120
ATKAVLQEFGRIDILVNNNG MSQRSLCMDT SLDVYRKIE LNYLGTVSLT KCVLPHMIER 180
KQGKIVTVNS ILGISVPLS IGYCASKHAL RGFFENGLRTE LATYPGIIVS NICPGPVQSN 240
15 IVENSLAGEV TKTIGNNGDQ SHKMTTSRCV RLMLISMAND LKEVWISEQP FLVTVYLWQY 300
MPTWAWWITN KMGKKRIENF KSGVDADSSY FKIFKTKHD

20

SEQ ID NO:141 PFH1 DNA SEQUENCE

Nucleic Acid Accession #: NM_021614
Coding sequence: 1-1740 (underlined sequences correspond to start and stop codons)

25

30

35

40

45

50

55

1 11 21 31 41 51
| | | | |
ATGAGCAGCT GCAGGTACAA CGGGGGCGTC ATGCGGCCGC TCAGCAACTT GAGCGCGTCC 60
CGCCGGAACC TGCACGAGAT GGAAGTCTG GCGCAGCCCC TGCAGCCCC CGCGTCTGTC 120
GGAGGAGGTG GCGGCGCGTC CTCCCCTCT GCAGCCGCTG CCGCCGCCGC CGCTGTTTCG 180
TCCTCAGCCC CCGAGATCGT GGTGTCTAAG CCCGAGCACA ACAACTCCAA CAACCTGGCG 240
CTCTATGGAA CCGGCGCGCG AGGCAGCACT GGAGGAGGCG GCGCGCGTGG CGGGAGCGGG 300
CACGGCAGCA GCAAGTGGC CAAAGTCCAGC AAAAAAGAAA ACCAGAACAT CGGCTACAAG 360
CTGGGCCACC GCGCGCCCT GTTCGAAAAG CGCAAGCGGC TCAGCGACTA CGCGCTCATC 420
TTCGGCATGT TCGGCATCGT GGTATGGTC ATCGAGACCG AGCTGTCTGT GGGCGCCTAC 480
GACAAGGCGT CGCTGTATT CTAGCTCTG AAATGCCTTA TCAGTCTCT CACGATCATC 540
CTGCTCGGTC TGATCATCGT GTACCACGCC AGGGAATAC AGTTGTTTAT GGTGGACAAT 600
GGAGCAGATG ACTGGAGAAT AGCCATGACT TATGAGCGTA TTTTCTTCAT CTGCTTGAA 660
ATACTGGTGT GTGCTATTCA TCCATACCT GGAATTATA CATTACATG GACGGCCCGG 720
CTTGCTTCT CATTGCCCC ATCCACAACC ACCGCTGATG TGGATATTAT TTTATCTATA 780
CCAATGTTCT TAAGACTCTA TCTGATTGCC AGAGTCATGC TTTACATAG CAAACTTTTC 840
ACTGATGCCT CCTCTAGAAG CATTGGAGCA CTTAATAAGA TAAACTTCAA TACACGTTTT 900
GTTATGAAGA CTTAATGAC TATATGCCA GGAAGTGTAC TCTTGGTTTT TAGTATCTCA 960
TTATGGATAA TIGCCGCATG GACTGTCCGA GCTTGTGAAA GGTACCATGA TCAACAGGAT 1020
GTTACTAGCA ACTTCCCTGG AGCGATGTGG TTGATATCAA TAACTTTTCT CTCCATTGGT 1080
TATGTGACA TGGTACCTAA CACATACTGT GGAAAAGGAG TCTGCTTACT TACTGGAATT 1140
ATGGGTGCTG GTTGCACAGC CTTGGTGGTA GCTGTAGTGG CAAGGAAGCT AGAAGTTACC 1200
AAAGCAGAAA AACACGTGCA CAATTTTCATG ATGGATACT AGCTGACTAA AAGAGTAAAA 1260
AATGCAGCTG CCAATGATG CAGGGAACA TGGCTAATT ACAAATAAC AAAGCTAGTG 1320
AAAAAGATAG ATCATGCAA AGTAAGAAA CATCAACGAA AATTCCTGCA AGCTATTCAT 1380
CAATTAAGAA GTGTAAAAAT GGAGCAGAGG AACTGAATG ACCAAGCAAA CACTTGGTG 1440
GACTTGGCAA AGACCCAGAA CATCATGTAT GATATGATTT CTGACTTAAA CGAAAGGAGT 1500
GAAGACTTCG AGAAGAGGAT TGTACCCCTG GAAACAAAAC TAGAGACTTT GATTGGTAGC 1560
ATCCACGCCC TCCCTGGGCT CATAAGCCAG ACCATCAGGC AGCAGCAGAG AGATTTTATT 1620
GAGGCTCAGA TGGAGAGCTA CGACAAGCAC GTCATTACA ATGCTGAGCG GTCCCGGTCC 1680
TCGTCCAGGA GCGGCGGTCT CTCTCCACA GCACCACCA CTTCATCAGA GAGTAGCTAG

SEQ ID NO:142 PFH1 Protein sequence:
Protein Accession #: NP_067627

60

65

70

1 11 21 31 41 51
| | | | |
MSSCRYNGGV MRPLSNLSAS RRNLHEMDSE AQPLPPASV GGGGGASSPS AAAAAAAVS 60
SSAPEIVVSK PEHNNNNLA LYGTGGGGST GGGGGGGGSG HGSSSGTKSS KKKNQNIYK 120
LGHRRALFEK RKRLSDYALI FGMFGIVVMV IETELSWGAY DKASLYSLAL KCLISLSTII 180
LLGLIIVYHA REIQLFMVDN GADDWRIAMT YERIFFICLE ILVCAIHP PNYTFTWTAR 240
LAFSYAPSTT TADVDILSI PMFLRLYLIA RVMLLHSLKF TDASSRSIGA LNKINFNTFR 300
VMKTLMTICP GTVLLVFSIS LWIAAWTVR ACERYHDQDQ VTSNFLGAMW LISITFLSIG 360
YGDMPVNTYC GKGVCLLTGI MGAGCTALVV AVVARKLELT KAEKHVHNF MDTQLTKRVK 420
NAAANVLRET WLIYKNTKLK KKIDHAKVRK HQRKFLQAIH QLRSVKMEQR KLNDQANTLV 480
DLAKTQNMIMY DMISDLNERS EDFEKRIVTL ETKLETIGS IHALPGLISQ TIRQQQRDFI 540
EAQMESYDKH VTYNAERSRS SSRRRSSST APPTSSESS

75

SEQ ID NO:143 PFG9 DNA SEQUENCE

Nucleic Acid Accession #: AL110139, coding region is FGENESH predicted
Coding sequence: 1-1896 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51

5
 10
 15
 20
 25
 30
 35

| | | | |
 ATGCGCGCCG TGCCGCTGCC CGCCCCGCTC CTGCCGCTGC TGCTGCTCGC GTCCTTGGCC 60
 GCTCCCGCCG CCGCGCCAG CAGAGCCGAG TCCGTCTCCG CGCCGTGGCC CGAACCCGAG 120
 CGCGAGTCGC GGCCACCGCC CGGCCGGGG CCCGGGAACA CCACCCGTT TGGGTCTGGG 180
 CGCGCGGGCG GCAGCGGCAG CTCCAGCTCC AACAGCAGTG GCGACGCCCTT GGTGACCCGC 240
 ATTTCCATCC TCCTCCGCGA CCTACCCACC CTCAAGGCAG CCGTGATCGT GGCCTTCGCC 300
 TTACCACCC TCCTCATCGC CTGCCCTGCTG CTGCGCGTCT TCAGTGCGGG AAAGAGGTTA 360
 AAGAAGACAC GCAAGTATGA TATCATCACC ACTCCAGCAG AGCGAGTGG AATGGCGCCA 420
 CTAAATGAAG AGGATGATGA AGATGAGGAC TCCACAGTAT TCGACATCAA ATACAGAGTG 480
 TCCTTGCCGG CTGCACTGAG ACCTCAGCTG CCAGGGTGCC AGACGCTACT GACAGTTCCT 540
 GTGCCCCAC CCTTCATCCT CGACATTGAC CTTCACGCAA GATGCAGTGG AAGGCCTGAT 600
 GGTGGAATCA GACCTGGTAA AACCTGTTTC CCAGCCTGGT GGCATCCTGT GGAAGTGG 660
 TCAGCTGCAA CTTGGGGTGT GAAGGACTGG ACCTGGAAGC CTTCTGCGT CGGAGGTGTT 720
 GAAACCAAAA CGAACGTTAT GTATAAAACC CCAGCTCCAT CGTGCGTGTC AGGCATCTGC 780
 TCAGACTGTC ACTGGCAAGC TCGTTTCAC GTCACCACAA TGGAGTTGCT TCTGCCACCC 840
 TTTGGGCATC CCTTTAAAGT GCCCCTACT TCTACTCCCC ATGGTTTTCG ACAACTGCAG 900
 CTGAATCTCA TGGAAAAGCT GGATTCCTCT GCCTTACGCA GAAACACCCG GGCTCCATCT 960
 GCCAGGTGCT TGCCACTGGT CTGGGCAGAA ATGGCGGCTG CTGAAAGTGA CCTTCCAAAT 1020
 CCTTGGTGGC ACTTCAGGC CACAGGCTCT CCAATAAAAA CCCTTTACAC ACAACCATG 1080
 AGTACCTTGG GCTTGGATGT TTTCTGTTGT GCCGCGCAGC GGGGCACCTT TTGTGAAGAC 1140
 AGAGCAGTGA CTAAGGTTCT CCAGGTAGC TCTTTCTCCA AACAGCTGCG CTGGAAGCCA 1200
 GCCCTAGAGA GTGGGTTTCC CCATCATCTC AGGCTTCTCA GAGAGTGTC TCCGCTGAGC 1260
 ACCCATCTCT TCAGGTTGGC TCGTTCAGAT GCCGGGGGAC AAGCCAGCCT GACGGGGAGG 1320
 AGGGTGTTC GCGTCCGCG GCAGTCTCTG CATGGCGGAG GGTGAGCGGG TACCGCACT 1380
 TGCCTTTTGG TTTGAAGAT TCTGTGAGG CGCCATCCTC ACCTTGACCT CTTCTCAAAA 1440
 ATCTGTCTCC CCGTCTGTGC CCGTGAACAC CTACGGGAAG CCAAGAGAAG CTCAGTGACT 1500
 GTCCTTGCCT CATTGAGCA GAGCCACAA AAGGCAGCTG CTGCCACGG GGAGCCTGTC 1560
 AAACGAGGGC CCACTGGGCA ATTGACCAGA CACACATGCC CTGGCTGGGG GATCACACAT 1620
 GCGAACCCTG AGACAATCC AGATACCAA GGCCAGGAAG GCCACGCTGA GGATGTCAT 1680
 CACCTGGAG GAGACTTGA TGGGTGGCA AATTCTATT TGGAGGAAGA GGGTTTCCAG 1740
 GATGTCAGAT GCCAGAAGAT GGTCTGATG TCTGAGGAAG GGCCACCTAG TTTGACAGGA 1800
 TGTGAGAGGC TCACAGGTTT CCATCATCTC TCCAGCCATT CCAAGTCTG GTCCTTCCTT 1860
 TCCCCCGAC AGCCCTGTT TCTGTCCAGG CCTTGA

SEQ ID NO:144 PFG9 Protein sequence:

Protein Accession #: none available, FGENESH predicted

40
 45
 50

1 11 21 31 41 51
 | | | | |
 MRVPLPAPL LPLLLLALLA APAARASRAE SVSAPWPEPE RESRPPPGPG PGNTRFRSG 60
 AAGSGSSSS NSSGDALVTR ISILLRDLPT LKAAVIVAF FTLIIACLL LRVFRSGKRL 120
 KKTRKYDIIT TPAERVEMAP LNEEDEDDED STVFDIKYRV SLPAALRRQL PGCQTLLTVP 180
 VPPFFILDIID LPARCSRFPD GGIRPGKTCF PAWWHPVESW SAATWGVKDW TWKPSCVGGV 240
 ETKTNVMYKT PAPSCVSGIC SDCHWQARFH VTMELLPP FGHFPKVPPT STPHGFRQLQ 300
 LNLMEKLDSS ALRRNRAPS ARCLPLVLA MAAESDLPN PWWHFSATGS PIKTLTYQTM 360
 STLGLDVFCG AGQRGTFCE RAVTKVLQGS SFSKQLRWKP ALESGFPHHL RLLRECPPLS 420
 THPVRLARSD ARGQASLTGR RVFRRRQSL HGGGSAGTAT CLLVLKILLR RHPHLDLFYK 480
 ICLPCCAVEH LREAKRSSVT VLSFEQSPQ KAAAHAHGEPV KRGP SGQLTR HTCPGWGITH 540
 ANLQTIPTDQ GQEGFREDVT HPGGDLGDVA NFYLEEEGFQ DGRGCKMVL M SEEGPPSLTG 600
 CERLTGSHHF SSHSKSWSFL SPRQLFLSR P

SEQ ID NO:145 PFG6 DNA SEQUENCE

Nucleic Acid Accession #: NM_013427

Coding sequence: 875-3799 (underlined sequences correspond to start and stop codons)

60
 65
 70
 75

1 11 21 31 41 51
 | | | | |
 GGCTGGGCTG CGAATAGCGT GTTCTCTCC GGCGGAACAC ACACACCCGG CCTTGGGGCT 60
 GTCTCTGAA GCTCCCTCCT CCACGGAGAG CGCTGAGCGC CGCCGGGAAT TCCATCCCAC 120
 CGTGGGCACG CAGTCTTTGG AGGTCCCGGG CGCAGCACGC TCGGTGTCC CACACTGCAG 180
 CAAGACAGAG ACCCGCGGGG AACCTTGAGC TTGGAACAAC CCTTGAGCCT CTGCACTCGG 240
 AAGAGTGGGC GCAGCAGCCC AGCGGAGGCC AGGCGCGCAA CTTGGGGCGC CGGGGCAAGG 300
 AGAGAGTGCA GGGAGGCGCA GCTCAGGCGC CCGGCTCAGG AGCGGGAGGA AGTTCTCGCG 360
 GCGCCGGGAG CGCGGTGGAC GCGCCCTGGG CGCAGCCTTC TCCTTGCCCC 420
 TCGGACTGT CTTGGGGCGG CAAGGAGGAG CTGCTGGAG TCTTAGAGGC CATCCAGAGC 480
 CAGCGAGCAG GAGCGCTGCG TCTCCGCTC CAGCTAGGAA GGGGGAGTGG CGCTGGCAGG 540
 CTGGAGCTGG GAACCCAGCG AGCGCCTGAC CTTCTCTCT CTCTCTCTGA CCCTCTTCGC 600
 GTCTTGGGT CCGGAGGAAG GTTCTAGCGG CTGCAAGAGG TCCCAGACC CATTTCCTA 660
 GAAGGCTGGT GATGGATCTG CTGCTCTGCG CGCCGCCGGG GCACTTGGAG CGCACCGGCG 720
 GCGCGTGAGC TGGGCTTTGC TCTCCACGCG CCTGGGCAAA CCCCAGGCA GCCCGGCTG 780
 GCACCTTTGC CTGAGTCCCT TCGGTTTCCC GACCCAAAGC CACAGCGTC CAGGGAGGGA 840
 GGAGGAGGTG GTCCTCAGT GCAGCCCCCG CGAGATGTCC GCGCAGAGCC TGCTCCACAG 900
 CGTCTTCTCC TGTTCTCTCG CCGTCTCAAG TAGCGCGGCC TCGGCAAGG GCTTCTCAA 960
 GAGGAAGCTG CGCCAGACCC GCAGCCTGGA CCGGCGCTG ATCGGCGGCT GCGGGAGCGA 1020
 CGAGGCGGGC GCGGAGGGCA GTGCGCGGGG AGCCACGCGG GGCCGCTCT ACTCCCCATC 1080
 ACTCCAGCC GAGAGTCTCG GCCCTGCTT GCGCTCTCT TCCCGGGGTC CGCCCCCAG 1140
 GGCCACCAGG CTACCGCTCT CTGGACCTCT TTGCTCGTCC TTCTCCACAC CCAGCACCCC 1200

GCAGGAGAAG TCACCATCCG GCAGCTTTCA CTTTGACTAT GAGGTTCCCC TGGGTCGCGG 1260
 CGGCCCTCAAG AAGAGCATGG CCTGGGACCT GCCTTCTGTC CTGGCCGGGC CAGCCAGTAG 1320
 CCGAAGCGCT TCCAGCATCC TCTGTTTCATC CGGGGGAGGC CCCAATGGCA TCTTCGCTTC 1380
 5 TCCTAGGAGG TGGCTCCAGC AGAGGAAGTT CCAGTCCCCA CCCGACAGTC GCGGGCACCC 1440
 CTACGTCGTG TGAATAATCC AGGGTGATTT CACCTGGAAC AGCATGTCTAG GCCGCAAGTGT 1500
 GCGGCTGAGG TCAGTCCCCA TCCAGAGTCT CTCAGAGCTG GAGAGGGCCC GGCTGCAGGA 1560
 AGTGCCTTTT TATCAGTTGC AACAGGACTG TGACCTGAGC TGTCAGATCA CCATTCCCAA 1620
 AGATGGACAA AAGAGAAAGA AATCTTTAAG AAAGAAACTG GATTCACTAG GAAAGGAGAA 1680
 10 AAACAAAGAC AAAGAATTCA TCCACAGGC ATTTGGAATG CCCTTATCCC AAGTCATTGC 1740
 GAATGACAGG GCCTATAAAC TCAAGCAGGA CTGACAGAGG GACGAGCAGA AAGATGCATC 1800
 TGACTTTGTG GCTTCCTCC TCCATTGAG AAATAAAAGA CAAAACAAAG AACTCTCAAG 1860
 CAGTAACCTA TCTCTCAGCT CAACCTCAGA AACACCGAAT GAGTCAACGT CCCCACACAC 1920
 CCCGGAACCG GCTCCTCGGG CTAGGAGGAG GGGTGCCATG TCAGTGGATT CTATCACCAG 1980
 TCTTGATGAC AATCAGTCTC GACTACTAGA AGCTTTACAA CTTCCTTGC CTGCTGAGGC 2040
 15 TCAAAGTAAA AAGGAAAAAG CCAGAGATAA GAAACTCAGT CTGAATCCTA TTTACAGACA 2100
 GGTCCCTAGG CTGGTGGACA GCTGCTGTCA GCACCTAGAA AAACATGGCC TCCAGACAGT 2160
 GGGGATATTC CGAGTTGGAA GCTCAAAAAA GAGAGTGAGA CAATTACGTG AGGAATTGTA 2220
 CCGTGGGATT GATGTCTCTC TGGAGGAGGA GCACAGTGT CATGATGTGG CAGCCTTGCT 2280
 GAAAGAGTTC CTGAGGGACA TGCCAGACCC CCTTCTCACC AGGGAGCTGT ACACAGCTTT 2340
 20 CATCAACACT CTCTGTGTGG AGCCGAGGA ACAGCTGGGC ACCITGCAGC TCCTCATATA 2400
 CCTTCTACCT CCCTGCAACT GCGACACCCT CCACCGCTG CTACAGTTCC TCTCCATCGT 2460
 GGCCAGGCAT GCCGATGACA ACATCAGCAA AGATGGGCAA GAGGTCACTG GGAATAAAAT 2520
 GACATCTCTA AACTTAGCCA CCATATTTGG ACCCAACCTG CTGCACAAGC AGAAGTCATC 2580
 AGACAAAGAA TTCTCAGTTC AGAGTTCAGC CCGGGCTGAG GAGAGCACGG CCATCATCGC 2640
 25 TGTTGTGCAA AAGATGATTG AAAATTATGA AGCCCTGTTC ATGGTTCCCC CAGATCTCCA 2700
 GAACGAAGTG CTGATCAGCC GTTAGAGAC CGATCTGAT GTCGTGGACT ATTTACTCAG 2760
 AAGAAAGGCT TCCCAATCAT CAAGCCCTGA CATGCTGCAG TCGGAAGTTT CTTTTCCGT 2820
 GGGAGGGAGG CATTTCATCTA CAGACTCCAA CAAGGCCTCC AGCGGAGACA TCTCCCTTA 2880
 TGACAAACAC TCCCAGTGTG TGTCTGAGCG CTCCTGCTG GCTATGCAAG AGGACGCGGC 2940
 30 CCCGGGGGCG TCGGAGAAGC TTTACAGAGT GCCAGGGCAG TTTATGCTGG TGGGCCACTT 3000
 GTCGTCGTCA AAGTCAAGGG AAAGTCTCTC TGGACCAAGG CTGGGAAAG ATCTGTCAGA 3060
 GGAGCCTTTC GATATCTGGG GAACCTGGCA TTCAACATTA AAAAGCGGAT CCAAAGACCC 3120
 AGGAATGACA GGTTCTCTCT GAGACATTTT TGAAGCAGC TCCCTAAGAG CGGGGCCCTG 3180
 CTCCCTTTCT CAAGGGAACC TGTCCCCAAA TTGGCCTCGG TGCAGGGGA GCCCCGAGA 3240
 35 GCTGGACAGC GACACGCGAG GGGCTCGGAG GACTCAGGCC GCAGCCCCCG CGACGGAGGG 3300
 CAGGGCCCCAC CTTGCGGTGT CGCGCGCCTG CAGCAGCCCC CACGTCCAGG TGGCAGGGAA 3360
 AGCCGAGCGG CCCACGGCCA GGTGCGGAGCA GTACTTGACC CTGAGCGCGC CCCACGACCT 3420
 CAGCGAGAGT GAGCTGGATG TGGCCGGGCT GCAGAGCCGG GCCACACCTC AGTGCCAAAG 3480
 40 ACCCATGGG AGTGGGAGGG ATGACAAGCG GCCCCCGCTT CCATACCCGG GCCCAGGGAA 3540
 GCCCGCGGCA GCGCGAGCCT GATCCAGGG GCCCCCGGAA GGCCTGGAGA CACCCACGGA 3600
 CCAGGAGGCG CAAGCAGCCG AGCGAGAGCA GCAGGTACAG CAGAAAAAAC TGAGCAGCGC 3660
 CAACTCCCTG CCAGCGGGCG AGCAGGACAG TCCGCGCCTG GGGGACGCTG GCTGGCTCGA 3720
 CTGGCAGAGA GAGCGCTGGC AGATCTGGGA GCTCCTGTG ACCGACAACC CCGATGCCCT 3780
 45 GCCGAGAGCG CTGCTCTGAG CCCGACCCA GCCGAGCCCC CCTGCCCCG AGCCCCCGCG 3840
 CCTCCAGCCC AGGGGGGACC GTGGGTGGTG GCCACTGGCA CACTTAGTGT TCTTCTTCA 3900
 CACTTCTCAA AAGTGACACA AGAGAAATCC AGTTCACCTA CAGAGGTAGA GCACTCACGC 3960
 CCCCGCCATT GAGAATAAGG TTCCATTGCG TAGCCAGCCT TAGGAAAAAC AAACAGAAC 4020
 CAAACCATAG GGCAATGTCC AATCTAAAAA CGTCCCTCTT GGCTCTATAA TATAAGATAC 4080
 50 AACTCTGCT TGGTATAGCC TAACGTAAT TATGTGTCTT CGGTTTGGAC TATTGTGTAT 4140
 TCTGTAACAG ATTATGTATA ATCATATATG ATATATTCAC AAAGAGAAAA CAAAAGGAAC 4200
 TTTTAAAAAA AATACTACTT CACTTATATT AAGCAATGAG ATATACTAAA CAATGAGATT 4260
 CTATAGAATG TTTAGAAATG TGCACAAGCG GGTTTCTGTG CTTTGGCCAT AGCTTTATAA 4320
 CTGGGGATAA CCCTTCCTTC GATACAAAC ACTAACAAGA GGAAGCAGAA TATGAGAAGC 4380
 55 CATATTTTAA CATAGGAGTC AGATACAAAA AGAAAAATCA CTGAATGCTT TTAGATATTG 4440
 AATACGTTTT CAGGAAAAATG CTAATCTGA TAGATTACGA AATATATTTT TAGAAGTTGT 4500
 TTAGAAAGGA TTCAGTTAAC CAAACAAGAA AAAGGCAGTG CCTCACAAG AAATTAAGAA 4560
 GTTGTCGCTC CCACGTTACA TCAAATTCAG TTTTATATAG GCCATATATA ATATATATT 4620
 ATAATGTATA ATTTTATGT ATTTTTCAAA ACTACAACT GGAATCCAAC TATAAGTGT 4680
 TTAAGAATCT ACACAGAATA TTCAAATAT AGAACATGTT TTTCCCTTT GCCCCATAAT 4740
 60 CAGTATTTGC CAAATTACAT GCAATTCCTT AAAAACTAAA TCACATTGGT AAAAGGCCTA 4800
 CAGCTTTGTA CTTACATTGT GCCAAAGGCT GAGGAAATGT TTCTTTTCGA ATTTTATGT 4860
 GTATTGTAAG ATGTTCTACC GTACTTTAGT AGTTTGAAGT TTTCAAGTGC ATAACATTT 4920
 TIGACCAGCA GAAGGCGATA CGCTTCAGTA TTTTATGCAA TTTTTCCTT CTTGGAAGGG 4980
 65 AAAGTGTATT ATAAAAAAG ATTTTTCCTT TTTAAACAT GCTACTCTTA ATTTTCATGT 5040
 TGGTGATGAA ATTCCAGTG GTGTTTCTTA AGTTTCTATC TTGTGCCATG ATGAATAAAA 5100
 AGTTAAGCAA AAAAAAAAAA AAAAAAAAAA AAA

SEQ ID NO:146 PFG6 Protein sequence:
 Protein Accession #: NP_038286.1

1 11 21 31 41 51
 MSAQSLHVS FSCSSPASSS AASAKGFSKR KLRQTRSLDP ALIGCGSDE AGAEGSARGA 60
 75 TAGRLYSPSL PAESLGPRLA SSSRGPPRA TRLPPLGGLC SSFSTPSTPQ EKSPSGSFHF 120
 DYEVLGRGG LKMSMAWDLF SVLAGPASSR SASSILCSSG GGPNGIFASP RRWLQQRKFQ 180
 SPPDSRGHPY VVWKSEGDFT WNSMSGRSVR LRSVPIQSL ELERARLQEV PFYQLQDQCD 240
 LSCQITPKD GQKRKSLRK KLDLKGKEKN KDKEFIPQAF GMPLSQVIAN DRAYKLQDL 300
 QRDEQKQDASD FVASLLPFGN KRQNKELSSS NSSLSSTSET PNESTSPNTP EPAPRRRRRG 360

AMSVDSITDL DDNQSRLLA LQLSLPAEAQ SKKEKARDKK LSLNPIYRQV PRLVDSCCQH 420
 LEKHGLQTVG IFRVGSSKKR VRQLREEFDR GIDVSLEEEH SVHDAALLK EFLRDMPPDL 480
 LTRELYTAFI NTLLEPEEQ LGTLQLLIYL LPPCNCDTLH RLLQFLSIVA RHADDNISKD 540
 GQEVTVGNKMT SLNLTATIFGP NLLHKQKSSD KEFSVQSSAR AEESTAIIV VQKMIENYEA 600
 LFMVPPDLQN EVLISLLETQ PDVVDYLLRR KASQSSSPDM LQSEVSFSVG GRHSSTDSNK 660
 ASSGDISPYD NNSPVLSESR LLAMQEDAAP GGSEKLYRVP GQFMLVGHLS SSKSRESSPG 720
 PRLGKDLSEE PFDIWTGWH S TLKSGSKDPG MTGSSGDIFE SSSLRAGPCS LSQGNLSPNW 780
 PRWQGSAPL DSDTQGARRT QAAAPATEGR AHPAVSRACS TPHVQVAGKA ERPTARSEQY 840
 LTLGAHDL S ESELDVAGLQ SRATPQCQRH HGSGRDDKRP PPPYPGPKP AAAAAWIQGP 900
 PEGVETPTDQ GGQAAREEQ VTQKKLSSAN SLPAGEQDSP RLGDAWLDW QRERWQIWEL 960
 LSTDNDALP ETLV

SEQ ID NO:147 PFG4 DNA SEQUENCE

Nucleic Acid Accession #: NM_002202

Coding sequence: 240-1289 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 CCCCCGAGCC GCGCCGAGTC TGCCGCCGCC GCAGCGCCTC CGCTCCGCCA ACTCCGCCGG 60
 CTTAAATGG ACTCTAGAT CCGCGAGGGC GCGGCGCAGC CGAGCAGCGG CTCTTTCAGC 120
 ATTGGCAACC CCAGGGGCCA ATATTTCCCA CTTAGCCACA GCTCCAGCAT CCTCTCTGTG 180
 GGCTGTTTAC CAACTGTACA ACCACCATTT CACTGTGGAC ATTACTCCCT CTTACAGATA 240
 TGGGAGACAT GGGAGATCCA CCAAAAAAAAA AACGTCTGAT TTCCCTATGT GTTGGTTGCG 300
 GCAATCAGAT TCACGATCAG TATATTCTGA GGGTTTCTCC GGATTGGAA TGGCATGCGG 360
 CATGTTTGAA ATGTGCGGAG TGTAATCAGT ATTTGGACGA GAGCTGTACA TGCTTTGTGA 420
 GGGATGGGAA AACCTACTGT AAAAGAGATT ATATCAGGTT GTACGGGATC AAATGCGCCA 480
 AGTGCAGCAT CGGCTTCAGC AAGAACGACT TCGTGATGCG TGCCCGCTCC AAGGTGTATC 540
 ACATCGAGTG TTCCCGCTGT GTGGCCTGCA GCCGCCAGCT CATCCTGGG GACGAATTG 600
 CGCTTCGGGA GGACGGTCTC TTCTGCCGAG CAGACCACGA TGTGGTGGAG AGGGCCAGTC 660
 TAGGCGCTGG CGACCCGCTC AGTCCCTGTC ATCCAGCGCG GCCACTGCAA ATGGCAGCGG 720
 AGCCCATCTC CGCCAGGCAG CCAGCCCTGC GGCCCCACGT CCACAAGCAG CCGGAGAAGA 780
 CCACCCGGT GCGGACTGTG CTGAACGAGA AGCAGCTGCA CACCTTGGCG ACCTGCTACG 840
 CCGCAAACCC GCGGCCAGAT GCGCTCATGA AGGAGCAACT GGTAGAGATG ACGGGCCTCA 900
 GTCCCCGTGT GATCCGGGTC TGGTTTCAAA ACAAGCGGTG CAAGGACAAG AAGCGAAGCA 960
 TCATGATGAA GCAACTCCAG CAGCAGCAGC CCAATGACAA AACTAATATC CAGGGGATGA 1020
 CAGGAACCTC CATGTTGGCT GCCAGTCCAG AGAGACACGA CGGTGGCTTA CAGGCTAACC 1080
 CAGTGGAAGT ACAAAGTTAC CAGCCACCTT GGAAAGTACT GAGCGACTTC GCCTTGCAGA 1140
 GTGACATAGA TCAGCTGTCT TTTCAGCAAC TGGTCAATTT TTCAGAAGGA GGACCGGGCT 1200
 CTAATTCAC TGCGAGTAA GTAGCATCAA TGTCTCTCA ACTTCCAGAT ACACCTAACA 1260
 GCATGGTAGC CAGTCTATT GAGCATGAG GAACATTCAT TCTGTATTT TTTCCCTGT 1320
 TGGAGAAAGT GGGAAATTA AATGTGCAAC TCTGAAACAA AAGTATTTAA CGACCCAGTC 1380
 AATGAAACT GAATCAAGAA ATGAATGCTC CATGAAATGC ACGAAGTCTG TTTAATGAC 1440
 AAGGTGATAT GGTAGCAACA CTGTGAAGAC AATCATGGGA TTTTACTAGA ATTAACAAC 1500
 AAACAAACG CAAAACCCAG TATATGCTAT TCAATGATCT TAGAAGTACT GAAAAAATA 1560
 GACGTTTTTA AAACGTAGAG GATTTATATT CAAGGATCTC AAAGAAAGCA TTTTCATTTT 1620
 ACTGCACATC TAGAGAAAAA CAAAAATAGA AAATTTTCTA GTCCATCCTA ATCTGAATGG 1680
 TGCTGTTTCT ATATTGGTCA TTGCCTTGCC AAACAGGAGC TCAGCAAAA GCGCAGGAAG 1740
 AGAGACTGGC CTCCTTGCT GAAAGAGTCC TTTCAGGAAG GTGGAGCTGC ATTGGTTTGA 1800
 TATGTTTAAA TGTGACTTTA ACAAGGGGTT AATTGAAATC CTGGTCTCT TGCCCTGTCC 1860
 TGTAGCTGGT TTATTTTATA CTTTGCCCCC TCCCCACTTT TTTTGAGATC CATCTTTAT 1920
 CAAGAAGTCT GAAGCGACTA TAAAGGTTT TGAATTCAGA TTTAAAAACC AACTTATAAA 1980
 GCATTGCAAC AAGGTACCT CTATTTTGCC ACAAGCGTCT CGGGATTGTG TTTGACTTGT 2040
 GTCTGTCCAA GAACCTTTCC CCCAAAGATG TGTATAGTTA TTGGTTAAAA TGACTGTTT 2100
 CTCTCTCTAT GGAATAAAAA AGGAAAAAAA AAAGGAAACT TTTTGTGTTT GCTCTTGCAT 2160
 TGCAAAAAAT ATAAAGTAAT TTATTATTTA TTGTCGGAAG ACTTGCCACT TTTTCATGCA 2220
 TTTGACATTT TTTGTTTGT GAAGTGAAAA AAAAAGATAA AGGTTGTACG GTGGTCTTTG 2280
 AATTATATGT TTAATCTAT TGTTTTGTCT TTTTCTTAA ATATTATGTG AAATCAAAAGC 2340
 GCCATATGTA GAATTATATC TTCAGGACTA TTCTACTAAT AAACATTTGG CATAGAT

SEQ ID NO:148 PFG4 Protein sequence:

Protein Accession #: NP_002193.1

1 11 21 31 41 51
 MGDPPKKKRL ISLCVCGNQ IHDQYILRV S PDLEWHAACL KCAECNQYLD ESCTCFVRDG 60
 KTYCKRDYIR LYGIKCAKCS IGFSKNDFVM RARSKVYHIE CFRVACSRQ LIPQDEFALR 120
 EDGLFCRADH DVVERASLGA GDPLSPLHPA RPLQMAAEPI SARQPALRPH VHKQPEKTR 180
 VRTVLNEQL HTLRTCYAAN PRPDALMKEQ LVEMTGLSPR VIRVWFQNK RCKDKRSIMM 240
 KQLQQQQPND KTNIQGMTGT PMVAASPERH DGLQLANPVE VQSYQPPWKV LSDFALQSDI 300
 DQFAFQQLVN FSEGGPGSNS TGSEVASMSS QLPDTPNSMV ASPIEA

SEQ ID NO:149 PFG2 DNA SEQUENCE

Nucleic Acid Accession #: NM_001172

Coding sequence: 39-1103 (underlined sequences correspond to start and stop codons)

5 1 11 21 31 41 51
 | | | | |
 GCGGAGCTCT GCCTTGGAGA TTCTCAGTGC TCGGATCAT GTCCCTAAGG GGCAGCCTCT 60
 CGCGTCTCCT CCAGACGCGA GTGCATTCCA TCCTGAAGAA ATCCGTCCAC TCCGTGGCTG 120
 10 TGATAGGAGC CCGCTTCTCA CAAGGGCAGA AAAGAAAAGG AGTGGAGCAT GGTCCCGCTG 180
 CCATAAGAGA AGCTGGCTTG ATGAAAAGGC TCTCCAGTTT GGGCTGCCAC CTAAAAGACT 240
 TTGGAGATTG GAGTTTACT CCAGTCCCCA AAGATGATCT CTACAACAAC CTGATAGTGA 300
 ATCCACGCTC AGTGGGTCTT GCCAACCAGG AACTGGCTGA GGTGGTTAGC AGAGCTGTGT 360
 CAGATGGCTA CAGCTGTGTC AACTGGGAG GAGACCACAG CCTGGCAATC GGTACCATTA 420
 GTGGCCATGC CCGACACTGC CCAGACCTTT GTGTTGCTG GGTGATGCC CATGCTGACA 480
 15 TCAACACACC CCTTACCATT TCATCAGGAA ATCTCCATGG ACAGCCAGTT TCATTCTCC 540
 TCAGAGAAGT ACAGGATAAG GTACCACAAC TCCCAGGATT TTCTGGATC AAACCTTGTA 600
 TCCTTCTGC AAGTATGTG TATATTGGTC TGAGAGACGT GGACCTCCT GAACATTITA 660
 TTTAAAGAA CTATGATATC CAGTATTTT CCATGAGAGA TATTGATCGA CTGGTATCC 720
 AGAAGTTCAT GGAACGAACA TTGATCTGCG TGATTGGCAA GAGACAAAGA CCAATCCATT 780
 20 TGAGTTTGA TTTGACCTA TTTGACCTA CACTGGCTCC AGCCACAGGA ACTCCTGTTG 840
 TCGGGGGACT AACCTATCGA GAAGGCATGT ATATTGCTGA GGAAATACAC AATACAGGGT 900
 TGCTATCAGC ACTGGATCTT GTTGAAGTCA ATCCTCAGTT GGCCACCTCA GAGGAAGAGG 960
 CGAAGACTAC AGCTAACCTG GCAGTAGATG TGATTGCTTC AAGCTTTGGT CAGACAAGAG 1020
 AAGGAGGGCA TATTGTCTAT GACCAACTTC CTACTCCAG TTCACCAGAT GAATCAGAAA 1080
 25 ATCAAGCAG TGTGAGAATT TAGGAGACAC TGTGCACTGA CATGTTTCAC AACAGGCATT 1140
 CCAGAATTAT GAGGCATTGA GGGGATAGAT GAATACTAAA TGGTTGCTG GGTCAATACT 1200
 GCCTTAATGA GAACATTAC ACATTCTCAC AATTGTAAAG TTCCCTCTCT ATTTTGGTGA 1260
 CCAATACTAC TGTAATGTA TTTGGTTTTT TGCAGTTCAC AGGTATTAA TATGCTACAG 1320
 TACTATGTAA ATTTAAAGAA GTCATAAACA GCATTATTAT CCTTGGTATA TCATACTGGT 1380
 30 CTGTTGCTG TGTTCCTTC ACATTAAAGT GGTTTTTCAT CTTCCTCCC TCCTCCACA 1440
 GCCTGGCTAT ACAGTGCATC CTGAACCTGT CAGCCACAG CAGCAATATG CTTATTCTAT 1500
 CCACATCCCT AACATCATGC ATTCAACAAG TCAAAGTTCT GGTCCACAAA CCCTCCCTA 1560
 TAGAAGTTCA ATGGCTGCGA AAGAATTGT AGTAAACCAG GCCTCCAGG ATGGCGAGCT 1620
 CCAGTAAGAT GATAATGGAA AGCAGCAGCT TGTGGTTGT CACTCTACAA AGAGAAGCAA 1680
 35 AGTGGGGAGT AGTCAGAAGT TTGGATAACC TTCTTCTAA ACATTGGGG GTTAGACCTG 1740
 GGACCACGGC TGTATCTCT GAGGCTGTAT GTTGATCAC ACAGCCACT AGCAGGAAGT 1800
 ACTCATAAGG TTCTTTAGCT GTCACCTAGG GATAACACTG TCTACCTCAC AGAAATGTTA 1860
 AACTGAGACA ATAAAACCCA AAGCAT

SEQ ID NO:150 PFG2 Protein sequence:

Protein Accession #: NP_001163.1

45 1 11 21 31 41 51
 | | | | |
 MSLRGSLSRL LQTRVHSILK KSVHSVAVIG APFSQGQKRK GVEHGPAAIR EAGLMKRLSS 60
 LGCHLKDFGD LSFTVPVKDD LYNNLVNPR SVGLANQELA EVVSRVSDG YSCVTLGGDH 120
 SLAIGTISGH ARHCPDL CVV WVDADADINT PLTSSGNLH GQPVSFLLRE LQDKVPQLPG 180
 50 FSWIKPCISS ASIVYIGLRD VDPPEHFILK NYDIQYFSMR DIDRLGIQKV MERTFDLLIG 240
 KRQRPILHSF DIDAFDPTLA PATGTPVVG LTYREGMYIA EEIHNTGLLS ALDLVEVNPQ 300
 LATSEEEAKT TANLAVDVIA SFGQTREGG HIVYDQLPTP SSPDESENQA RVRI

SEQ ID NO:151 PFG1 DNA SEQUENCE

Nucleic Acid Accession #: NM_017906

Coding sequence: 80-1255 (underlined sequences correspond to start and stop codons)

60 1 11 21 31 41 51
 | | | | |
 AATTATATAT TTTTACTCTA TGTTTCTCTA CATGTTTTTT TCTTCCGTT GCTGGCGGAA 60
 GAGGCACGTG CGCTGCTGAA TGGAGCTGGT CGCTGGTTGC TACGAGCAGG TCCTCTTTGG 120
 GTTCGCTGTA CACCCGAGC CCAAGGCTTG CGGCGACCAC GAGCAATGGA CTCTTGTTGGC 180
 65 TGACTTCACT CACCATGCTC AACTGCCTC CTGTGTCAGCA GTAGCTGTAA ATAGTCGTTT 240
 TGTGGTCACT GGGAGCAAAG ATGAAACAAT TCACATTAT GACATGAAAA AGAAGATTGA 300
 GCATGGGGCT CTAGTGATC ACAGTGGTAC AATAACTTGC CTGAAATTCT ATGGCAACAG 360
 GCATTAAATC AGTGGAGCGG AAGATGGACT CATCTGTATC TGGGATGCAA AGAAATGGGA 420
 ATGCCTGAAG TCAATTAAG CTCACAAAGG ACAGGTGACC TTCCTTTCTA TTCACCCATC 480
 70 TGGCAAGTTG GCCCTGTCGG TTGGTACAGA TAAAACTTTA AGAACGTGGA ATCTTGTTAGA 540
 AGGAAGATCA GCATTCTATA AAAATATAAA ACAAAATGCT CACATAGTAG AATGGTCCCC 600
 AAGAGGAGAG CAGTATGTG TTATCATACA GAATAAAATA GACATCTATC AGCTTGACAC 660
 TGCATCCATT AGTGGACCA TCACAAATGA AAAGAGAATT TCCTCTGTTA AATTTCCTTC 720
 AGAGTCTGTC CTGTCAGTGG CTGGAGATGA AGAAGTTATA AGGTTTTTTG ACTGTGATTG 780
 75 ACTAGTGTGC CTCTGCGAAT TTAAGCTCA TGAACACAGG GTAAAGGACA TGTTCAGTTT 840
 TGAAATTCCA GAGCATCATG TTATTGTTTC AGCATCGAGT GATGGTTTCA TCAAAATGTG 900
 GAAGCTTAAG CAGGATAAGA AAGTTCCTCC ATCTTTACTC TGTGAAATAA ACTAATGC 960
 CAGGCTGACG TGTCTTGGAG TGTGGCTAGA CAAAGTGCCA GACATGAAAA GCCTTCTCTC 1020
 AGCTGCAGAG CCTTCTCTG TAAGTAAAGA ACAGTCCAAA ATTGGCAAAA AGGAGCTTGG 1080
 TGACACAGTG CACAAAGAAG AAAAGCGGTC AAAACCTAAC ACAAGAAAC GCGGTTTAAAC 1140

AGGTGACAGT AAGAAAGCAA CAAAAGAAAG TGGCCTGATA TCAACCAAGA AGAGGAAAAAT 1200
 GGTAGAAATG TTGGAAAAAG AGAGGAAAAA GAAGAAAAATA AAAACAATGC AGTGAATCAC 1260
 AGATGTCTCC TGAAGAAGCT CTTTATAGATG AAATCATTCT ACTCAAATGT ACCTTAATTT 1320
 TTTTITTTCC CTGAGTAAAA GCAAGAAATT TCTTCCTTTG GAAAAAATAT ATATATTAAA 1380
 AAACCACTTT TAGATGGTTT TTTTAAAAA AAAAAAAAAA ACTGGTAAAA TTACTTTTGG 1440
 CAGACAGTGT TTTATGAATT ATGTATCATG TTGATATATA ATATGTTAAT GTGTCATGTA 1500
 ATTTTACTT TGTACAAAGC AAATAAGAT CTTTCTCAAA AAAAAAAAAA AAAA

SEQ ID NO:152 PFG1 Protein sequence:
 Protein Accession #: NP_060376.1

1 11 21 31 41 51
 | | | | |
 MELVAGCYEQ VLFQFAVHPE PKACGDHEQW TLVADFTHHA HTASLSAVAV NSRFVVTGSK 60
 DETIHIVDMK KKHIEGALVH HSGTITCLKF YGNRHLISGA EDGLICIWDA KKWECLKSIK 120
 AHKGQVTFLS IHPSGKLALS VGTDKTLRTW NLVEGRSAFI KNKQNAHIV EWSRGEQYV 180
 VIIQNKIDYI QLDIASISGT ITNEKRISV KFLSESVLAV AGDEEVIRFF DCDLSVCLCE 240
 FKAHENRVKD MFSFEIPEHH VIVSASSDGF IKMWKLKQDK KVPSLLCEI NTNARLTCLG 300
 VWLDKVDAMK SLPPAAEPSP VSKEQSKIGK KEPGDTVHKE EKRSKPNTKK RGLTGDSKKA 360
 TKESGLISTK KRKMVEMLEK KRKKKKIKTM Q

SEQ ID NO:153 PFD6 DNA SEQUENCE

Nucleic Acid Accession #: NM_014668
 Coding sequence: 110-2953 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 GATGCTTGG ACATGCTCTG GCTGGCTAAT CTCCATGTTT TAGCCGACTG AAAATACGGT 60
 GGCCAAGTGG ATGGTGTGCT TATTTCAGT CTAAAGAAAT TTCTTTTGA TGTGGCAGAA 120
 AATCGAGGAT GTGGAGTGGG GACCCAGAGC TTAATTGGAG CTGGAGGGTC TGCCTTGCAT 180
 CTGATCTTC AGTGGGATGG ACCCGCATGG GGAGTCCTTG CCGAGGTCTT TGAGGTACTG 240
 TGACCTGCGA TTGATAAACT CCTCTGCTT GGTGAGAACA GCCTTGGAGC AGGAGCTGGG 300
 CCTGGCTGCC TACTTTGTGA GCAACGAGGT TCCCTTGGAG AAGGGGGCTA GGAACGAGGC 360
 CTTGGAGAGT GATGCTGAGA AGCTGAGCAG CACAGACAAC GAGGATGAGG AGCTGGGGAC 420
 AGAAGGCTCT ACCTCGGAGA AGAGAAGCCC CATGAAAAGG GAGAGGTCCC GCTCCACGA 480
 CTCAGCATCC TCATCCCTCT CCTCCAAGGC TTCCGGTTCA GCGCTCGGTG GCGAGTCCTC 540
 GGCTCAGCCC ACAGCACTCC CCCAGGGAGA GCATGCCAGG TCGCCCCAGC CCCGTGGCCC 600
 CGCAGAGGAG GGCAGAGCCC CTGGTGAGAA ACAGAGGCCC CGGGCAAGTC AGGGGCCACC 660
 GTCGGCCATC AGCAGGCACA GTCCCGGGCC GACGCCCCAG CCCGACTGTA GCCTCAGGAC 720
 CGGCCAGAGG AGCTTCCAGG TGTCGGTTCAC CTCGTCGTGC TCCAGCTGT CCTCTCTCTC 780
 GGGCTCATCC TGCTCATCCG TGGCGCCCGC TGCCGGCAGC TGGGTCTGCG AGGCCTCCCA 840
 GTGCTCCTTG ACCAAGGCCT GCCGCCAGCC ACCCATTTGTC TTCTTGCCCA AGCTCGTGTA 900
 CGACATGGTT GTGTCCACTG ACAGCAGTGG CTTGCCCAAG GCCGCTCCC TCCTGCCCTC 960
 CCCCTCGGTC ATGTGGGCGA GCTCTTCCG CCCCCTGCTC AGCAAGACCA TGACATCCAC 1020
 CGAGCAGTCC CTCTACTACC GGCAGTGGAC GGTGCCCCGG CCCAGCCACA TGGACTACGG 1080
 CAACCGGGCC GAGGGCCCGG TGGACGGCTT CCACCCCGC AGGCTGCTGC TCAGCGGCC 1140
 CCCTCAGATC GGAAGAGACG GTGCCCTACCT GCAGTTCTC AGTGTCCTGT CCAGGATGCT 1200
 TGTTGCGGTC ACAGAAGTGG ATGTCTATGA CGAGGAGGAG ATCAATATCA ACCTCAGAGA 1260
 AGAATCTGAC TGGCAATTAT TCCAGCTTAG CGACCCCTGG CCAGACCTGG AGCTGTTCAA 1320
 GAAGTTGCCC TTTGACTACA TCATTCACGA CCCGAAGTAT GAAGATGCCA GCCTGATTG 1380
 TTCGCACTAT CAGGGTATAA AGAGTGAAGA CAGAGGGATG TCCCGGAAGC CGGAGGACCT 1440
 TTATGTGCGG CGTCAGACGG CACGGATGAG ACTGTCCAAG TACGCAGCGT ACAACACTTA 1500
 CCACCACTGT GAGCAGTGCC ACCAGTACAT GGGCTTCCAC CCCCCTACC AGCTGTATGA 1560
 GTCCACCCTG CACGCCTTG CTTCTCTTA CTCCATGCTA GGAGAGGAGA TCCAGCTGCA 1620
 CTTTCATATC CCCAAGTCCA AGGAGCACCA CTTTGTCTT AGCCAACCTG GAGGCCAGCT 1680
 GGAGAGCATG CGATACCCCC TCGTGACAGA CAAGAGCCAT GAATATATAA AAAGTCCGAC 1740
 ATTCATCCA ACCACCGGCC GTCACGAACA TGGGCTCTT AATCTGTACC ACGCAATGGA 1800
 CGGTGCCAGC CATTTGCACG TGCTGGTTGT CAAGGAATAC GAGATGGCAA TTTATAAGAA 1860
 ATATTGGCCC AACCACATCA TGCTGGTGCT CCCAGTATC TTCAACAGTG CTGGAGTTGG 1920
 TGCTGCTCAT TTCTCATCA AGGAGCTGTC CTACCATAA CTGGAGCTCG AGCGGAACCG 1980
 GCAGGAGGAG CTGGGAATCA AGCCGCAGGA CATCTGCCT TTCAATTGTA TCTCTGATGA 2040
 CTCCTGCGTG ATGTGGAACG TGGTGGATGT CAACTCTGCT GGGGAGAGAA GCAGGGAGTT 2100
 CTTCTGGTGC GAAAGGAACG TGTCTTTGAA GCACATCATG CAGCACATCG AGCGGGCCCC 2160
 CGACATCATG CACTACGCC TGCTGGGCT GCGGAAGTGG TCCAGCAAGA CCGGGGCCAG 2220
 CGAGGTGCAA GAGCCCTCT CCGCTGCCA CGTGACAAC TTCATCATCC TGAACGTGGA 2280
 CCTGACCCAG AACGTGCAGT ACAACCAGAA CCGGTTCTCT TGTGACGATG TAGACTTCAA 2340
 CCTGCGGGTG CACAGCGCCG GCCTCTGCT CTGCGGTTT AACCGCTTCA GCGTGATGAA 2400
 GAAGCAGATC TGGGTGGGCG GCCACAGGTC CTTCCACATC ACATCCAAGG TGTCTGATAA 2460
 CTCTGCCGCG GTCTGCGCG CCCAGTACAT CTGTGCCCG GACAGCAAGC ACACGTTCTT 2520
 CGCAGCGCCC GCCAGCTCC TGCTGGAGAA GTTCTCTGAG CACCACAGCC ACCTCTTCTT 2580
 CCGCTGTCC CTGAAGAAACC ATGACCACCC AGTGCTGTCT GTCGACTGTT ACCTGAACCT 2640
 GGGATCTCAG ATTCTGTTT GCTATGTGAG CTCCAGGCC CACTCTTAA ACATCAGCTG 2700
 CTCGGACTTG CTGTTCAAGT GGCTGCTGCT GTACCTCTGT GACTCTTTT TGGGAGCTAG 2760
 CTTTTTGAAG AAGTTTCATT TTCTGAAAGG TGCGACGTTG TGTGTCATCT GTCAGGACCG 2820
 GAGCTCACTG CGCCAGACGG TCGTCCGCT GGAGCTCGAG GACGAGTGGC AGTTCCGGCT 2880

GCGCGATGAG TTCCAGACCG CCAATGCCAG GGAAGACCGG CCGCTCTTTT TTCTGACGGG 2940
 ACGACACATC TGAGGAAGAC AGCGGCGAGT TTTCTGAAGA GATGAGTGCT CAGAGCCCTC 3000
 ATGCTGTGTA GGCTAAAGGG AGGCCTGGAA CGGTGGGGCG TTGACTGGA ATGACCCCA 3060
 5 GGGACTGTCC AGGTGCAGCC CCTCCTAGTA CACATGGGCC CCCGAGGCCG TGGTCCTGGG 3120
 AGCCAGGAAG ACTCCGCAGT GGGTGAGAAT GAAACTTGA GACTCCCAAG TTCTGGGCCA 3180
 GCCCATTGCT CTGGGCTGTT TTAAGGCCCA TTTCACGAGG AACAAAGATT TACTTCCTGT 3240
 CCTGCCATTG GTGTGCTTCC ATGGACAAAC CTGATTTTTT TCTCTTAGTT CTAAAGAATC 3300
 TTGGGTATT TTGTAGCGGT GCCAGTATT CAGTAGATGG GATTCAGCC AAGTAGGTTT 3360
 10 CCCTGTAACC TCCTACAAAG CAATATTCCA AAGGAACATT TTAAGTGTA AGGCTGGAGA 3420
 CAAGAAAAAA TAAGTAGATC GTTTTAATAA CAATTATTTA ATTGCCTATA AGTTTGCTGT 3480
 TTCAGAGGCT AGCCCAAAGG CATCAAATTT AATAAAGTTA AACAAATTGA TTAACCTCAG 3540
 AGCAAATATG ATCCTATTAA AATAATATAG GGTAAATACC CTACCTCTTA GAAAGGGCAA 3600
 AAATGCAAAG AAGCTTTCCT TAAACTAAA AGGGTTTTTT GGGGGGGGAG TTGGCGGGGA 3660
 15 GGAATAAAGG CTAACAGAGG TTGACCTAAA ATTAGCCTTA CAAAGGAGAA AGGACCACAT 3720
 TGCTTACTTG AAACAGACAA TGAAACAAC CAAAGTGATA TATAAATAG TTGATGAGAA 3780
 CTAGACTTAT GACTGTAGTT TACTAGAGTT TAGTTTTTCTG TGCTGAAGT AGCTCATTTT 3840
 CTCTTACTAA TGTTTGGTTC CTCAGGGAAG AATCTCACTT GACTAGAGAG GAGGTGGGAA 3900
 CAGAAGAGAG AAGGAGGCAG GGAGATGTAT TTCTTAGGGC TCACCCCTTC ACAGACTGAC 3960
 AGAATGGTTT TGTTTTGTGT TGTTTTGTGT TGTTTTGTGT TTGAGATGGA CTCTAGCTCT 4020
 20 GTCACCCAGG CTGGAGTGCA GTGGTGCGAT CTCGGCTCAC TGCAAGCTCC GCCTCCCGGG 4080
 TTCTACCAT TCTCTGCTC CAGCCTCCCG AGTAGCTGGG ACTACAGCGC CCCACCACCA 4140
 CGCCCGGCTA ATTTTGTGTA TTTTITAGTA GAGACGGGGT TTCACCATGT TAGCCAGGAT 4200
 GGTCGCGATC TCCTGACCTC GTGATCCGCC CGCCTCGGCC TCCCAAAGTG CTGGGATTAC 4260
 25 AGGCGTGAGC CACCGTGCTT GCCCAGAAT GGTTTTTTAA GCCACAGTTG AGAGGCCACC 4320
 CATTGCCCGG CGCTCGGACA GTGATCATCT TGTTCATCTT GTTCAGTCTT TTCTGTGTGT 4380
 ATTGGAATTA TATCTCCCTT TTGAAAGATG AGAAGGTTGA GATGCAAAGA GTCTACCTTT 4440
 CCAAGTTCTC ACTGCTGGAA AGAGCTAGAA GCACAGTTCA AAGTCTGGC TTCTGGACTC 4500
 TGCAGTCCAG GTCTCCCTTC TCCCACTGCT CTACCCCTCA TGCCACACTG TTTTGAAGT 4560
 GGCCCAATC TTGAAGGAAA AGTTTAAAGA CAGTCAATT TAATCATCAG AATGCATCTT 4620
 30 TTTTITTTTC GGAGACGGAG TTTCACCTCT GCTGCCAGG CTGGAGTGCA ATGGTGCAAT 4680
 GATCTCGGCT CACTGCAACC TCTGCCTCT GGGTTCAAGT GATTCTCCAG CCTCAGCCTC 4740
 CCGAGTAGCT GGGATTATGG GCGCCACCA CCATGCCAG CTAATTTTGT TATTTTITTT 4800
 TTTTAGTAGA GATGGGGTTC CGCCAGGTTG GCCAGGCTGG TCTGTGAAC TCCTGGCCTC 4860
 AGGTGATCTG CCCACCTCAT CTCTCCAAAAG TGCTGGGATT ACAGGCATGA GCCACTGCGC 4920
 35 CTGGCCTCAG AATGCATTCT TACACATCTA TCCTAGACAT TTATAAGCAC TCTAATGGAT 4980
 AACAAATCAA GAATAAATGA TTGTAAAAGA TGATGCCGAA GAGTTGATGT CAATCTTTTT 5040
 TTCTAAGAA AAAAAGTCCG CGAGTATTA ATATTTAGAT CAATGTTTAT AAAATGATTA 5100
 CTTTGTATAT CTCATTATTC CTATTTTGA ATAAAACTG ACCTCTCTTA ATCATATACT 5160
 40 TGTCTTTTGT AAATAGCAGC TTTTGTGTC TTTCCCCAC TTTATTAGTT AATTAAAT 5220
 GGAAAAAACC CTCAACTAA TATTCTGTG TGTTCAGTC TTATAAATAA AACTTATAAT 5280
 GCATG

45 SEQ ID NO:154 PFD6 Protein sequence:
 Protein Accession #: NP_055483.1

1 11 21 31 41 51
 | | | | |
 50 MWQKIEDVEW RPQTYLELEG LPCILFSGM DPHGESLPRS LRYCDLRLIN SSCLVRTALE 60
 QELGLAAAYFV SNEVPLEKGA RNEALESDAE KLSSTDNEDE ELGTEGSTSE KRSPMKRERS 120
 RSHDSASSSL SSKASGALG GESSAQPTAL PQGEHARSPO PRGPAEEGRA PGEKQRPRAS 180
 QGPPSAISRH SPGPTPQPDCL SLRTGQRSVQ VSVTSSCSQL SSSSGSSSSS VAPAAAGTWVL 240
 55 QASQCSLTKA CRQPIVFLP LLYVDMVST DSSGLPKAAS LLPSPVMWA SFRPLLSKT 300
 MTSTEQSLYY RQWTVPRPSH MDYGNRAEGR VDGPHRRLL LSGPPQIGKT GAYLQFLSVL 360
 SRMLVRLTEV DVEDEEINI NLREESDWHY LQLSDPWPDL ELFKKLPFDY IHDPKYEDA 420
 SLICSHYQGI KSEDGRMSRK PEDLYVRRQT ARMRLSKYAA YNTYHCEQC HQYMGFHPRY 480
 QLYESTLHAF AFSYSMLGEE IQLHFIPKS KEHFFVFSQP GGQLESMLRP LVTDKSHEYI 540
 60 KSPTFTPTTG RHEHGLFNLY HAMDGASHLH VLVVKEYEMA IYKKYWPNNH MLVLPISFNS 600
 AGVGAHFLI KELSYHNLEL ERNRQEELGI KPQDIWPFIV ISDDSCVMWN VVDVNSAGER 660
 SREFSWERN VSLKHIMQHI EAAPDIMHYA LLGLRKWSSK TRASEVQEPF SRCHVHNFI 720
 LNVDLTQNVQ YNQNRFLCDD VDFNLRVHSA GLLLCRFNRF SVMKKQIVVG GHRSFHITSK 780
 VSDNSAAVVP AQYICAPDSK HTFLAAPAQL LLEKFLQHHS HLFPLSLKN HDHPVLSVDC 840
 65 YLNLSGSQSV CYVSSRPDSL NISCDLLES GLLLYLCSDF VGASFLKKFH FLKGATLCVI 900
 CQDRSSLRQT VVRLELEDEW QFRLRDEFQT ANAREDRPLF FLTGRHI

SEQ ID NO:155 PFC6 DNA SEQUENCE

70 Nucleic Acid Accession #: NM_000522
 Coding sequence: 1-1167 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 75 ATGACAGCCT CCGTGCTCCT CCACCCCGCG TGGATCGAGC CCACCGTCAT GTTCTCTAC 60
 GACAACGGCG CGGCCTGGT GGCCGACGAG CTCAACAAGA ACATGGAAGG GCGGCGGCG 120
 GCTGACGACG CGGCTGCAGC GCGGCGGCTT GCCGGGGCCG GGGGCGGGGG CTCCCCAC 180
 CCGGCGGCTG CCGCGGCGAG GGGCAACTTC TCGGTGGCGG CCGCGGCGCG GGCTGCGGCG 240
 GCGGCGGCGG CGCAACGATG CCGCAACCTG ATGGCGCACC CGGCGCCCTT GCGCCAGGA 300
 GCCGCGTCCG CCTACAGCAG CGCCCCGGG GAGGCGCCCC CGTCGGCTGC GCGGCTGCT 360

GCCGCGGCTG CCGCTGCAGC CGCCGCCGCC GCCGCGCGCT CGTCTCGGG AGGTCCCGGC 420
 CCGGCGGGCC CGGCGCGGCC AGAGGCGGCC AAGCAATGCA GCCCTGCTC GGCAGCGGCG 480
 CAGAGCTCGT CGGGGCCCGC GCGCTGCC TATGGCTACT TCGGCAGCGG CTACTACCG 540
 TCGCGCCGCA TGGGCCCGCC CCCAACGCC ATCAAGTCGT GCCCCAGCC CCCCTCGGCC 600
 GCGCGCCCGC CCGCTTTCG GACAAAGTAC ATGGATACCG CCGGCCAGC TGCCGAGGAG 660
 TTCAGTCCC GCGCTAAGGA GTTCGCGTTC TACCACCAGG GCTACGCAGC CGGGCCTTAC 720
 CACCACCATC AGCCCATGCC TGGTACCTG GATATGCCAG TGGTCCGGG CCTCGGGGGC 780
 CCGGCGAGT CGCGCCACGA ACCCTTGGGT CTTCCCATGG AAAGTACCA GCCTTGGGCG 840
 CTGCCAACG GCTGGAACGG CCAATGTAC TGCCCCAAG AGCAGGCGCA GCCTCCCCAC 900
 CTCTGGAAGT CACTCTGCC CGACGTGGT TCCCATCCCT CGGATGCCAG CTCCTATAG 960
 AGGGGGAGAA AGAAGCGCGT GCCTTATACC AAGGTGCAAT TAAAGAAGT TGAACGGGAA 1020
 TACGCCACGA ATAAATTCAT TACTAAGGAC AAACGGAGGC GGATATCAGC CACGACGAAT 1080
 CTCTCTGAGC GGCAGGTCAC AATCTGGTTC CAGAACAGGA GGGTTAAAGA GAAAAAAGT 1140
 ATCAACAAAC TGAACACAC TAGTTAA

SEQ ID NO:156 PFC6 Protein sequence:
 Protein Accession #: NP_000513.1

1 11 21 31 41 51
 MTASVLLHPR WIEPTVMFLY DNGGLVADE LKNMEGAAA AAAAAAAAAA AGAGGGGEPH 60
 PAAAAAGGNF SVAAAAAAA AAAANQCRNL MAHPAPLAPG AASAYSSAPG EAPPSAAAAA 120
 AAAAAAAGAA AAASSSGGPG PAGPAAAEAA KQCSPCSAAA QSSSGPAALP YGYFGSGYYP 180
 CARMGPPNNA IKSCPQPSA AAAAAFADKY MDTAGPAAEE FSSRAKEFAF YHQGYAAGPY 240
 HHHQPMGYL DMPVVPGLGG PGESRHEPLG LPMESYQFWA LPNGWNGQMY CPKEQAQPPH 300
 LWKSTLPDVG SHPSDASSYR RGRKRVPYT KVQLKELERE YATNKFITKD KRRRISATTN 360
 LSERQVTIWF QNRRVKEKKV INKLKTTT

SEQ ID NO:157 PFA3 DNA SEQUENCE

Nucleic Acid Accession #: AW102723
 Coding sequence: 523-2676 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 CCCTTATGGC GATTGGGCGG CTGCAGAGAC CAGGACTCAG TTTCCCTGCC CTAGTCTGAG 60
 CCTAGTGGGT GGGACTCAGC TCAGAGTCAG TTTTCAGAA GAGGTTTCAG TTGCAGAGTT 120
 TTCCTACACT TTTCCTGCGC TAGAGCAGCG AGCAGCCTGG AACAGACCCA GCGCGAGGAC 180
 ACCGTGGGGG GAGGAGGCGC CTGGAGGAGC TTAGAGACCC CAGCCGGGCG TGATCTCACC 240
 ATGTGCGGAT TTGCGAGGCG CGCCTGGAG CTGCTAGAGA TCCGGAAGCA CAGCCCGCAG 300
 GTGTGCGAAG CCACCAAGAG TGCGGCTCTT GGAGAAAGCG TGAGCAGGGG GCCACCGCGG 360
 TCTCCGGCCT GTCTGCACCC TGTCGCTGAG GCTGCCTGAC AGTGACAATG ACATCCCACT 420
 TACCAGTGTC CTGAATTGA TAGTGGCTTC TGTGTGTCAG TCTCATATAA GAACTACAGC 480
 TCATCAGGAG GAGATCGCAG CAGGGTAAGA GACACCAACA CATTGTTCTG CACGAAGCTC 540
 AAGGATCTCA AGATCAGAGG AGAGTGTCTT TTTCTTAC TGGCACCAGG TCAAGTTCTT 600
 AACGAGTCTT CAGAGGAGGC AGCAGGAAGC TCAGAGAGCT GCAAAGCAAC CGTGCCCATC 660
 TGTCAAGACA TTCTTGAGAA GAACATACAA GAAAGTCTTC TCAAAAGAAA AACCAAGTCG 720
 AGCCGAGTCT ATCTTCACAC TTTGGCAGAG AGTATTTGCA AACTGATTTT CCCAGAGTTT 780
 GAACGGCTGA ATGTGTGACT TCAGAGAACA TTGGCAAAGC ACAAAATAAA AGAAAGCAGG 840
 AAATCTTTGG AAGAGGAAGA CTTTGAAGAA ACAATTGCA AGCAAGCAGT GCAGCAGAGT 900
 CCAGTGGAGT TATCAAGAAA TCTCTGGGTG AAGAGGTTTT TAAATATGT TACGAGGAAG 960
 ATGAAACACT CCTTGGGGTG GTTGGAGGCA CCCTTAAAGA TTTTAAACA GCTTCAGTAC 1020
 CCTTCTGAAA CAGAGCAGCC ATTGCCAAGA AGCAGGAAAA AGGGGACAGT TGAGGACGCC 1080
 TCCATTCTAT GCCTGGATAA GGAGGATGAT TTTCTACATG TTTACTACTT CTTCCTTAAG 1140
 AGAACCACCT CCTGATTCTT TCCCGGCATC ATAAAGGCAG CTGCTCACGT ATTATATGAA 1200
 ACGGAAGTGG AAGTGTCTGT AATGCCTCCC TGCTTCATA ATGATTGCAG CGAGTTTGTG 1260
 AATCAGCCCT ACTGTGTGTA CTCCGTTTAC ATGAAAAGCA CCAAGCCATC CCTGTCCCCC 1320
 AGCAAAACCC AGTCTCTGCT GGTGATTCCC ACATCGCTAT TCTGCAAGAC ATTTCCATTG 1380
 CATTTCATGT TTGACAAAGA TATGACAATT CTGCAATTG GCAATGGCAT CAGAAGGCTG 1440
 ATGAACAGGA GAGACTTTCA AGGAAAGCCT AATTTTGAAT ACTTTGAAAT TCTGACTCCA 1500
 AAAATCAACC AGACCTTTAG CGGGATCATG ACTATGTTGA ATATGCAGTT TGTGTACGA 1560
 GTGAGGAGAT GGGACAACCT TGTGAAGAAA TCTTCAAGGG TTATGGACCT CAAAGGCCAA 1620
 ATGATCTACA TTGTTGAATC CAGTGCAATC TTGTTTTTGG GGTCAACCTG TGTGGACAGA 1680
 TTAGAAGATT TTACAGGACG AGGGCTCTAC CTCTCAGACA TCCCAATTCA CAATGCACCTG 1740
 AGGGATGTGG TCTTAATAGG GGAACAAGCC CGAGCTCAAG ATGGCCTGAA GAAGAGGCTG 1800
 GGGAAAGCTGA AGGCTACCTT TGAGCAAGCC CACCAAGCCC TGGAGGAGGA GAAGAAAAAG 1860
 ACAGTAGACC TCTGTGCTC CATATTTCCC TGTGAGGTTG CTCAGCAGCT GTGGCAAGGG 1920
 CAAGTTGTGC AAGCCAAGAA GTTCAGTAAT GTCACCATGC TCTTCTCAGA CATCGTTGGG 1980
 TTCACCTGCA TGTGCTCCA GTGCTCACCG CTGCAGGTCA TCACCATGCT CAATGCACCTG 2040
 TACACTCGCT TCAGCEAGCA GTGTGGAGAG CTGGATGTCT ACAAGGTGGA GACCAATTGCG 2100
 ATGCCTATTG TGTGGCTTGG GGGATTACAC AAAGAGAGTG ATACTCATGC TGTTTCAGATA 2160
 GCGCTGATGG CCCTGAAGAT GATGGAGCTC TCTGATGAAG TTATGTCTCC CCATGGAGAA 2220
 CCTATCAAGA TGCGAATTGG ACTGCACCTT GGATCAGTTT TTGCTGGCGT CGTTGGAGTT 2280
 AAAATGCCCC GTTACTGTCT TTTTGGAAAC AATGTCACTC TGGCTAACAA ATTTGAGTCC 2340
 TGCAGTGTAC CACGAAAAAT CAATGTCAAG CCAACAACCT ACAGATTACT CAAAGACTGT 2400
 CCTGGTTTCG TGTTTACCCC TCGATCAAGG GAGGAACCTC CACCAAACTT CCCTAGTGAA 2460
 ATCCCGGAAA CTTCGCAATT TCTGGATGCT TACCAACAAG GAACAACTC AAAACCATGC 2520

TTCCAAAAGA AAGATGTGGA AGATGCAAGC CAATTTTTTA GGCAAAGCAT CAGGAATAGA 2580
 TTAGCAACCT ATATACCTAT TTATAAGTCT TTGGGGTTTG ACTCATTGAA GATGTGTAGA 2640
 GCCTCTGAAA GCACCTTTAGG GATTGTAGAT GGCTAACAAG CAGTATTAAA ATTTCAGGAG 2700
 CCAAGTCACA ATCTTTCTCC TGTTTAACAT GACAAAATGT ACTCACTTCA GTACTTCAGC 2760
 TCTTCAAGAA AAAAAAAAAA ACCTTAAAAA GCTACTTTTG TGGGAGTATT TCTATTATAT 2820
 AACCAGCACT TACTACCTGT ACTCAAAATT CAGCACCTTG TACATATATC AGATAATTGT 2880
 AGTCAATTGT ACAAACCTGAT GGAGTCACCT GCAATCTCAT ATCCTGGTGG AATGCCATGG 2940
 TTATTAAAGT GTGTTTGTGA TAGTTGTCGT CAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 3000
 AAAA

SEQ ID NO:158 PFA3 Protein sequence:
 Protein Accession #: NP_000847.1

1 11 21 31 41 51
 MFCTKLKDLK ITGECFSL APGQVPLESS EEAAGSSESC KATVPICQDI PEKNIQESLP 60
 QRKTSRSRVY LHTLAESICK LIFEFERLN VALQRTLAKH KIKESRKSLE REDFEKTIAE 120
 QAVQQSPVEL SKNLLVKRFL KYVTRKMKTS LGWLEAPLKI FKQLQYPSSET EQPLPSRKK 180
 GQLEDASILC LDKEDDFLHV YYFFPKRTTS LILPGHIKAA AHVLYETEVE VSLMPPCFHN 240
 DCSEFVNQPY LLYSVHMKST KPSLSFSKPQ SSLVIPTSLF CKTFPFHFMF DKDMTILQFG 300
 NGIRRLMNRD DFGQKPNFEY FEILTPKINQ TFSGIMTMLN MQFVVRVRRW DNSVKKSSRV 360
 MDLKGQMIYI VESSAILFLG SPCVDRLEDF TGRGLYLSDI PIHNALRDVV LIGEQARAQD 420
 GLKKRLGKLK ATLEQAQHAL EEEKKKTVDL LCSIFPCEVA QQLWQQQVQV AKKFSNVTML 480
 FSDIVGFTAI CSQCSPLQVI TMLNLYTRF DQCCGELDVY KVETIAMPV WLGGLHKESD 540
 THAVQIALMA LKMMELSDV MSPHGEPIKM RIGLHSGSVF AGVVGVKMPR YCLFGNNVTI 600
 ANKFESCSVP RKINVSPTTY RLLKDCPGFV FTPRSREELP PNFPSEIPGI CHFLDAYQQG 660
 TNSKPCFQKK DVEDASQFFR QSRNRLATY IPIYKSLGFD SLKMCRASES TLGIVDG

SEQ ID NO:159 PFA1 DNA SEQUENCE

Nucleic Acid Accession #: NM_004362
 Coding sequence: 102-1934 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 CGCCGGCGGG ACTGGTCTGA AGAGACGCGG GGACAAAGTG GCAACGACTT GGACATCTGA 60
 GCTGTCACTG CCGAAAACAG GCCGCAAGAG AGATAATCAA TATGCATTTC CAAGCCTTTT 120
 GGCTATGTTT GGGTCTTCTG TFCATCTCAA TTAATGCAGA ATTTATGGAT GATGATGTTG 180
 AGACGGAAGA CTTTGAAGAA AATTCAGAAG AAATTGATGT TAATGAAAGT GAACCTTCCT 240
 CAGAGATTAA ATATAAGACA CCTCAACCTA TAGGAGAAGT ATATTTTGCA GAAACTTTTG 300
 ATAGTGAAG GTTGGCTGGA TGGGTCTTAT CAAAAGCAAA GAAAGATGAC ATGGATGAGG 360
 AAATTTCAAT ATACGATGGA AGATGGGAAA TTGAAGAGTT GAAAGAAAAA CAGGTACCTG 420
 GTGACAGAGG ACTGGTATTA AAATCTAGAG CAAAGCATCA TGCAATATCT GCTGTATTAG 480
 CAAAACCATT CATTTTGTCT GATAAACCTT TGATAGTTCA ATATGAAGTA AATTTTCAAG 540
 ATGGTATTGA TTGTGGAGGT GCATACATTA AACTCCTAGC AGACACTGAT GATTGTGATT 600
 TGGAAAACTT TTATGATAAA ACATCCTATA TCATTATGTT TGGACCAGAT AAATGTGGAG 660
 AAGATTATAA ACTTCATTTT ATCTTCAGAC ATAAACATCC CAAAACATGA GTTTTCGAAG 720
 AGAAACATGC CAAACCTCCA GATGTAGACC TTAATAAAGT CTTTACAGAC AGGAAGACTC 780
 ATCTTTATAC CTTTGTGATG AATCCAGATG ACACATTTGA GGTGTTAGTT GATCAAACAG 840
 TTGTAACAAA AGGAAGCCTC CTAGAGGATG TGGTTCCTCC TATCAAAACCT CCAAAGAAA 900
 TTGAAGATCC CAATGATAAA AAACCTGAGG AATGGGATGA AAGAGCAAAA ATCTCTGATC 960
 CTCTGCGGT CAAACAGAAA GACTGGGATG AAAGTGAACC TGCCCAATA GAAGATTCAA 1020
 GTGTGTGTTAA ACCTGCTGGC TGGCTTGATG ATGAACCAAA ATTTATCCCT GATCCTAATG 1080
 CTGAAAAAACC TGATGACTGG AATGAAGACA CGGATGGAGA ATGGGAGGCA CTCAGATTG 1140
 TTAATCCAGC ATGTCCGATT GGGTGTGGTG AGTGGAACCC TCCATGATA GATAACCCAA 1200
 AATACAAAGG AGTATGGAGA CCTCCACTGG TCGATAATCC TAACATCAG GGAATCTGGA 1260
 GTCCTCGAAA AATTCCTAAT CCAGATTATT TCGAAGATGA TCATCCATTT CTCTGACTT 1320
 CTTCAGTGC TCTTGGTTTA GAGCTTGGT CTATGACCTC TGATATCTAC TTTGATAATT 1380
 TTATTATCTG TTCCGAAAAA GAAGTAGCAG ATCACTGGGC TGCAGATGGT TGGAGATGGA 1440
 AAATAATGAT AGCAAAATGCT AATAAGCCTG GTGTATTAAA ACAGTTAATG GCAGCTGCTG 1500
 AAGGGCACCC ATGGCTTTGG TTGATTTATC TTGTGACAGC AGGAGTGCCA ATAGCATTA 1560
 TTACTTCATT TTGTGGGCA AGAAAAGTAA AGAAAAACA TAAAGATACA GAGTATAAAA 1620
 AAACCGACAT ATGTATACCA CAAACAAAAG GAGTACTAGA GCAAGAAGAA AAGGAAGAGA 1680
 AAGCAGCCCT GAAAAACCA ATGGACCTGG AAGAGGAAAA AAAGCAAAAT GATGGTGAAA 1740
 TGCTTGAAAA AGAAGAGGAA AGTGAACCTG AGGAAAAGAG TGAAGAAGAA ATTGAAATCA 1800
 TAGAAGGGCA AGAAGAAAGT AATCAATCAA ATAAGTCTGG GTCAGAGGAT GAGATGAAAG 1860
 AAGCAGATGA GAGCACAGGA TCTGGAGATG GGCCGATAAA GTCAGTACGC AAAAGAAGAG 1920
 TACGAAAGGA CTAACATAGA TTGAAATATT TTTAATTCCC GAGAGGATGT TTGGCATTGT 1980
 AAAAAACAGC ATGCCAGACC TGAACCTTAA TCAGTCTGCA CATCTGTTT CTAATATCTA 2040
 GCAACATTAT ATCTTTTCAG ACATTTATTT TAGTCCTTCA TTTCCGAGGA AAAAGAAGCA 2100
 ACTTTGAAGT TACCTCATCT TTGAATTTAG AATAAAAGTG GCACATTACA TATCGGATCT 2160
 AAGAGATTAA TACCATTAGA AGTTACACAG TTTTAGTTGT TTGGAGATAG TTTTGGTTTG 2220
 TACAGAACAA AATAATATGT AGCAGCTTCA TTGCTATTGG AAAAATCAGT TATTGGAATT 2280
 TCCACTTAAA TGGCTATACA ACAATATAAC TGGTAGTTCT ATAATAAAAA TGAGCATATG 2340
 TTCTGTGTG AAGAGCTAAA TGCAATAAAG TTTCTGTATG GTTGTGTGAT TCTATCAACA 2400
 ATTGAAAGTG TTGATATGA CCAACATTTA CCTAGTTTGT GTCAAAATTAT AGTTACAGTG 2460
 AGTTGTTTGC TTAATTTATA GATTCTTTTA AGGACATGCC TTGTTCAATA AATCACTGGA 2520

TTATATTGCA GCATATTTTA CATTGAATA CAAGGATAAT GGGTTTTATC AAAACAAAAT 2580
 GATGTACAGA TTTTTTTTCA AGTTTTTATA GTTGCTTTAT GCCAGAGTGG TTTACCCCAT 2640
 TCACAAAATT TCTTATGCAT ACATTGCTAT TGAAAATAAA ATTAAATAT TTTTTCATCC 2700
 TGAIAAAAAA

SEQ ID NO:160 PFA1 Protein sequence:
 Protein Accession #: NP_004353.1

1 11 21 31 41 51
 | | | | |
 MHFQAFWLCL GLLFISINAE FMDDDVETED FEENSEEIDV NESELSSEIK YKTPQPIGEV 60
 YFAETFDSEGR LAGWVLSKAK KDDMDDEEISI YDGRWEIEEL KENQVPGDRG LVLKSRAXHH 120
 AISAVLAKPF IFADKPLIVQ YEVNFQDGDID CGGAYIKLLA DTDDLLENF YDKTSYIMF 180
 GPDKCGEDYK LHFIFRHKHP KTGVFEEKHA KPPDVLKKF FTRKTHLYT LVMNPDFTFE 240
 VLVDQTVVVK GSLLDQVVP IKPKKEIEDP NDKKPEEWDE RAKIPDPSAV KPEDWDESEP 300
 AQIEDSSVVK PAGWLDDEPK FIPDPNAEKP DDWNEDTDGE WEAPQLNPA CRIGCGEWKP 360
 PMIDNPKYKG VWRPPLVDNP NYQGIWSPRK IPNPDYFEDD HPFLTSFSA LGLELWSMTS 420
 DIYFDNFIC SEKEVADHWA ADGWRWKIMI ANANKPGVLK QLMAAAEGHP WLWLIYLVTA 480
 GVPIALTSF CWPRKVKKKH KDTEYKTDI CIPQTKGVL EEEKEEKAAL EKPMDLSEEEK 540
 KQNDGEMLEK EEESEPEEKS EEEIEIEGQ EESNQSNKSG SEDEMKEADE STGSGDGPPIK 600
 SVRKRVRKRD

SEQ ID NO:161 PEZ9 DNA SEQUENCE

Nucleic Acid Accession #: NM_005932
 Coding sequence: 75-2216 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 GCGGAGCGCG CGCTCCAGC GAAAGCAGCA GGGCAGGGAT CTGCGTTGGA GGAAGGGACT 60
 GCTCTGGTGC TAGAATGCTG TGCCTCGGAA GGCTGGGCGG CTTGGGAGCC AGAGCAGCAG 120
 CTCTGCCGCC CCGCCGGGCG GGCCGGGGAA GCCTCGAAGC CGGGATCCGG GCCCGAAGGG 180
 TCAGCACCAG CTGGTCTCCC GTGGGCGCCG CCTCAATGT CAAGCCCCAG GGCAGCCGCT 240
 TGGACCTGTT CGGCGAGCGG GCGCGTCTTT TTGGAGTTCC TGAGCTGAGT GCCCCAGAAG 300
 GATTTCATAT TGCACAAGAA AAAGCCTTGA GAAAGACAGA ATTGCTTGTTG GACCGTGCAT 360
 GTTCCACCCC ACCTGGGCCC CAGACCGTGC TGATCTTCGA TGAGCTCTCG GATTCCCTAT 420
 GCAGAGTGGC GCAGCTGGCT GATTTGTGA AAATCGTCA CCCTGAGCCA GCATTCAAGAG 480
 AAGCTGCGGA AGAAGCTTGT AGAAGTATTG GCACCATGGT AGAGAAGTTG AACACAAATG 540
 TGGATTATA TCAAAAGTTT CAAAAATTAC TAGCTGATAA AAAACTTGTG GATTCCCTTG 600
 ATCCAGAAAC AAGCGAGTG CTGAAGTGT TATGTTTGA TTTTGAAAT AGTGGAATCC 660
 ATCTAGACAA ACAAAGCGT AAAAGAGCAG TGGACCTCAA TGTTAAAATC TTGGATTGTA 720
 GTAGTACATT TCTTATGGGA ACCAATTTTC CCAACAAGAT TGAGAAGCAT CTCTTACCAG 780
 AACACATTCT TCGTAACCTT ACATCTGCTG GGGATCATAT CATAATTGAT GGTCTCCACG 840
 CAGAATCACC AGATGACTTG GTGCGAGAAG CTGCTTATAA AATTTTCTT TATCCCAATG 900
 CTGGTCAATT GAAATGTTT GAAGAATTGC TCAGCAGCAG AGATCTTCTG GCAAAGTTGG 960
 TGGGGTATTC CACGTTTCT CACAGGGGTC TCCAAGGAAC GATAGCTAAA AATCCAGAGA 1020
 CTGTATGCA GTTCCTTGAA AAACATCTCT ACAAACCTTC TGAAGAAGCT CTGAAAGATT 1080
 TTGAGATGAT ACGAGGGATG AAAATGAAAC TGAATGCTCA AAATTCCGAA GTAATGCCCT 1140
 GGGACCCCCC TTAATACAGT GGTGTGATTC GTGCAGAAAG GTATAATATT GAGCCAGGCC 1200
 TATATGCCC GTTTTCTCT CTGGGAGCAT GCATGGAAGG CCTGAATATT TTGCTTAAAC 1260
 GACTGTGGG GATTCTTATA TATGCAGAGC AGCCTGCAAA AGGAGAGGTG TGGAGCGAAG 1320
 ATGTCCGAAA ACTGGCTGTT GTTCATGAAT CTGAAGGATT GTTGGGGTAC ATTTACTGTG 1380
 ATTTTITTC ACGAGCAGAC AAACCACATC AGGATTGCCA TTCTACTATC CGTGGAGGCA 1440
 GACTAAAGGA AGTGGAGAG TATCAACTCC CACTGTAGT TCTTATGCTG AATCTTCCCC 1500
 GTTCTCAAG GAGTCTTCCA ACTTTGCTAA CTCCTGGCAT GATGGAAAAAT CTTTCCATG 1560
 AAATGGGACA TGCCATGCAT TCAATGCTAG GACGTACTCG TTACCAACAC GTCACTGGGA 1620
 CCAGGTGCCC TACTGATTTT GCTGAGGTTT CTCTATTCT GATGGAGTAC TTGCAAAATG 1680
 ATTATCGAGT AGTTAACCAG TTTGCCAGAC ATTATCAGAC TGGACAGCCA CTGCCAAAAA 1740
 ATATGGTGTG TCGTCTTGT GAATCTAAAA AGGTTTGTGC TGCAGCTGAT ATGCAACTTC 1800
 AGGTCTTTTA TGCCACTCTG GATCAATCT ACCATGGGAA GCATCCCCTG AGGAATTCAA 1860
 CCACAGACAT TCTCAAGGAA ACACAAGAGA AATTCTATGG CCTACCATAT GTTCCAAATA 1920
 TCGCTGGCA GCTGCGATT AGCCACCTCG TGGGGTATGG TGCTAGATAT TACTCTTACC 1980
 TCATGTCAG AGCGGTGCGC TCCATGGTTT GGAAGGAGTG TTTTCTACAG GATCCTTTCA 2040
 ACAGGGCTGC CGGGGAGCGC TATCGCAGGG AGATGCTGGC CCACGGTGA GGCAGGGAGC 2100
 CCATGCTCAT GGTGAAGGT ATGCTTCAGA AGTGTCTTC TGTGTATGAC TTCGTAAGTG 2160
 CCCTCGTTT CGACTTGGAT CTGGACTTCG AAACCTTCTC CATGGATTCT GAATAAAGA 2220
 AACACTTAC ACCTCTAATC AAGGTATGT AGTAATGACT TTGTTATAAA TGCTACAGCT 2280
 GTGAGAGCTT GTTCTGATT GTTTCATTGT TCGTTCTGT AATTCTGAAA AACTTTAAAC 2340
 TGGTAGAACT TGAATAAAT AATTTGTTTT AATTAATAAA AAAAAAAAAA AA

SEQ ID NO:162 PEZ9 Protein sequence:
 Protein Accession #: NP_005923.1

1 11 21 31 41 51
 | | | | |
 MLCVGRLLGL GARAAALPPR RAGRSLEAG IRARRVTSW SPVGAAFNVK PQGSRLDLFG 60
 ERARLFGVPE LSAPEGFHIA QEALRKTEL LVDRACSTPP GPQTVLIFDE LSDSLCRVAD 120

LADFVKIAHP EPAFREAAEE ACRSIGTMVE KLTNTVDLYQ SLQKLLADKK LVDSLDPETR 180
 RVAELFMDFD EISGHLDDKQ KRKRAVDLNV KILDLSSTFL MGTNFPNKIE KHLLEPHIRR 240
 NFTSAGDHII IDGLHAESPD DLVREAAKYI FLYPNAGQLK CLEELLSSRD LLAKLVGYST 300
 FSHRALQGTI AKNPETVMQF LEKLSDDLSE RILKDFEMIR GMKMKLNAQN SEVMPWDPPY 360
 YSGVIRAERY NIEPSLYCFP FSLGACMEGL NILNRLLLGI SLYAEQPAKG EVWSEDVRKL 420
 AVVHESEGLL GYIYCDFFQR ADKPHQDCHF TIRGGRLKED GDYQLPLVVL MLNLPSSRS 480
 SPTLLTPGMM ENLFHEMGAH MHSMLGRTRY QHVTGTRCPT DFAEVPISLM EYFANDYRVV 540
 NQFARHYQTG QPLPKNMVSR LCESKKVCAA ADMQLQVFYA TLDQIYHGKH PLRNSTTDIL 600
 KETQEKFYGL PYVPNTAWQL RFSLVGYGA RYYSYLSMSRA VASMVWKECF LQDPFNRAAG 660
 ERYRREMLAH GGGREPMLMV EGMLQKCPV DDFVSALVSD LDLDFFETFLM DSE

SEQ ID NO:163 PEZ8 DNA SEQUENCE

Nucleic Acid Accession #: AF103907

Coding sequence: none (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 20 ACAGAAGAAA TAGCAAGTGC CGAGAAGCTG GCATCAGAAA AACAGAGGGG AGATTGTGT 60
 GGCTGCAGCC GAGGAGACC AGGAAGATCT GCATGGTGGG AAGGACCTGA TGATACAGAG 120
 GAATTACAAC ACATATACCT AGTGTTCCTAA TGAACACCAA GATAAATAAG TGAAGAGCTA 180
 GTCCGCTGTG AGTCTCCTCA GTGACACAGG GCTGGATCAC CATCGACGGC ACTTCTGTAG 240
 25 TACTCAGTGC AGCAAAAGAAA GACTACAGAC ATCTCAATGG CAGGGGTGAG AAATAAGAAA 300
 GGCTGCTGAC TTTACCATCT GAGGCCACAC ATCTGCTGAA ATGGAGATAA TTAACATCAC 360
 TAGAAACAGC AAGATGACAA TATAATGTCT AAGTAGTGAC ATGTTTTTGC ACATTTCCAG 420
 CCCCTTTAAA TATCCACACA CACAGGAAGC ACAAAGGAA GCACAGAGAT CCCTGGGAGA 480
 AATGCCCGGC CGGCATCTTG GGTATCATGAT GAGCCTCGCC CTGTGCCTGG TCCCGCTTGT 540
 30 GAGGAAGGA GATTAAAAA TGAATTGATG TGTTCCTTAA AGGATGGGCA GGAACACAGA 600
 TCCTGTGTGG CATATTATTT TGAACGGGAT TACAGATTGG AAATGAAGTC ACAAAGTGAG 660
 CATTACCAAT GAGAGGAAAA CAGACGAGAA AATCTTGATG GCTTCACAAG ACATGCAACA 720
 AACAAAAATG AATCATGTGA TGACATGAGG CAGCCAAGCT GGGGAGGAGA TAACCACGGG 780
 GCAGAGGGTC AGGATTCTGG CCCTGCTGCC TAACTGTGTC GTTCATAACC AAATCATTTC 840
 35 ATATTCTAA CCTCAAAAC AAAGCTGTG TAATATCTGA TCCTACGGT TCCTTCTGGG 900
 CCCAACATTC TCCATATATC CAGCCACACT CATTTTAAAT ATTTAGTICC CAGATCTGTA 960
 CTGTGACCTT TCTACACTGT AGAATAACAT TACTCATTTT GTTCAAAGAC CCTTCGTGTT 1020
 GCTGCCTAAT ATGTAGCTGA CTGTTTTTCC TAAGGAGTGT TCTGGCCAG GGGATCTGTG 1080
 AACAGGCTGG GAAGCATCTC AAGATCTTTC CAGGGTTATA CTTACTAGCA CACAGCATGA 1140
 40 TCATTACGGA GTGAATTATC TAATCAACAT CATCCTCAGT GTCTTTGCC ATACTGAAAT 1200
 TCATTCCCA CTTTGTGACC CATCTCAAG ACCTCAAAAT GTCAATCCAT TAATATCACA 1260
 GGATTAACCT TTTTTTTTAA CCTGGAAGAA TTCAATGTTA CATGCAGCTA TGGGAATTTA 1320
 ATTACATATT TTGTTTCTCA GTGCAAGAT GACTAAGTCC TTATCCCTC CCCTTTGTTT 1380
 GATTTTTTTT CCAGTATAAA GTTAAATGTC TTAGCCTTGT ACTGAGGCTG TATACAGCAC 1440
 45 AGCCTCTCCC CATCCCTCCA GCCTATCTGT TCATCACCAT CAACCCCTCC CATACCACCT 1500
 AAACAAATC TAACCTGTAA TTCCTGAAC ATGTCAGGAC ATACATTATT CCTTCTGCCT 1560
 GAGAAGCTCT TCCTGTCTC TTAATCTAG AATGATGTAA AGTTTGAAT AAGTTGACTA 1620
 TCTTACTTCA TGCAAGAAGG GGACACATAT GAGATTCATC ATCAGATGAG ACAGCAAATA 1680
 CTAAGAAGTG ATTTGTGTTA TAAGAGTTTA GATAAATATA TGAATGCAA GAGCCACAGA 1740
 50 GGAATGTTT ATGGGGCAGC TTTGTAAGCC TGGGATGTGA AGCAAAGGCA GGAACCTCA 1800
 TAGTATCTTA TATAATATAC TTCATTCTC TATCTCTATC ACAATATCCA ACAAGCTTTT 1860
 CACGAAATTC ATGCAGTGCA AATCCCAAAA GGTAACCTTT ATCCATTTC TGGTGAGTGC 1920
 GCTTTAGAAT TTTGGCAAAT CATACTGGTC ACTTATCTCA ACTTTGAGAT GTGTTGTGCC 1980
 TTGTAGTTAA TTGAAAGAAA TAGGGCACTC TTGTGAGCCA CTTAGGGTT CACTCCTGGC 2040
 55 AATAAAGAA TTACAAAGAG CTACTCAGGA CCAGTTGTTA AGAGCTCTGT GTGTGTGTGT 2100
 GTGTGTGTGT GAGTGTATAT GCCAAAGTGT GCCTCTCTCT CTGACCCAT TATTTCAGAC 2160
 TAAAAACAAG CATGTTTTC AATGGCACTA TGAGCTGCCA ATGATGTATC ACCACCATAT 2220
 CTCATTATTC TCCAGTAAAT GTGATAATAA TGTCTCTGT TAACATAAAA AAAGTTTGAC 2280
 TTCACAAAAG CAGCTGGAAG TGGACAACCA CAATATGCAT AAATCTAACT CCTACCATCA 2340
 60 GCTACACACT GCTTGACATA TATTGTTAGA AGCACCTCGC ATTTGTGGGT TCTCTTAAAG 2400
 AAAATACTTG CATTAGGTCT CAGCTGGGGC TGTGCATCAG GCGGTTTGAG AAATATTCAA 2460
 TTCTCAGCAG AAGCCAGAAT TTGAATTCCT TCATCTTTTA GGAATCATTT ACCAGGTTTG 2520
 GAGAGGATTC AGACAGCTCA GGTGCTTTCA CTAATGTCTC TGAATCTCTG TCCCTCTTTG 2580
 TGTTCATGGA TAGTCCAATA AATAATGTTA TCTTTGAAC TATGCTCATA GGAGAGAATA 2640
 65 TAAGAAGTCT GAGTGATATC AACATTAGGG ATTCAAAGAA ATATTAGATT TAAGCTCACA 2700
 CTGGTCAAAA GGAACCAAGA TACAAAGAAC TCTGAGCTGT CATCGTCCCC ATCTCTGTGA 2760
 GCCACAACCA ACAGCAGGAC CCAACGCATG TCTGAGATCC TTAATCAAG GAAACCAAGT 2820
 TCATGAGTTG AATCTCTCTA TTATGGATGC TAGCTTCTGG CCATCTCTGG CTCTCTCTTT 2880
 GACACATATT AGTCTCTAGC CTTTGTCTCC ACGACTTTTA TCTTTCTCC AACACATCGC 2940
 70 TTACCAATCC TCTCTCTGCT CTGTTGCTTT GGACTTCCCC ACAAGAATTT CAACGACTCT 3000
 CAAGTCTTTT CTCCATCTCC CACCCTAAC CTGAATGCCT AGACCCTTAT TTTTATTAAT 3060
 TTCCAAATAG TGCTGCTAT GGGCTATATT GCTTTAGATG AACATTAGAT ATTTAAAGCT 3120
 CAAAGGTTTC AAAATCCAAC TCATTATCTT CTCTTCTTT CACCTCCCTG CTCTCTCC 3180
 TATATTACTG ATTGCACTGA ACAGCATGGT CCCCAATGTA GCCATGCAAA TGAGAAACCC 3240
 75 AATGGCTCCT TGTGGTACAT GCATGCAAGA CTGCTGAAGC CAGAAGGATG ACTGATTACG 3300
 CCGTATGGGT GGAGGGGACC ACTCCTGGGC CTTCTGTGAT GTCAGGAGCA AGACCTGAGA 3360
 TGCTCCTGTC CTTAGTGTCT CTCTGCATCT CCCCTTTCTA ATGAAGATCC ATAGAATTTG 3420
 CTACATTTGA GAATTTCAAT TAGGAATCTA CATGTTTAT CTGCCCTATC AATTTTTTAA 3480
 ACTTGCTGAA AATTAAGTTT TTCAAAATC TGTCTTGTGA AATTACTTTT TCTTACAGTG 3540
 TCTTGGCATA CTATATCAAC TTTGATTCTT TGTACAACCT TTTCTTACTC TTTTATCACC 3600

AAAGTGGCTT TTATTCTCTT TATTATTATT ATTTTCTTTT ACTACTATAT TACGTTGTTA 3660
 TTATTTTGTG CTCTATAGTA TCAATTTATT TGATTTAGTT TCAATTTATT TTTATTGCTG 3720
 ACTTTTAAAA TAAGTGATTC GGGGGGTGGG AGAACAGGGG AGGGAGAGCA TTAGGACAAA 3780
 TACCTAATGC ATGTGGGACT TAAAACCTAG ATGATGGGTT GATAGGTGCA GCAAACCACT 3840
 ATGGCACACG TATACCTGTG TAACAAACCT ACACATTCTG CACATGTATC CCAGAACGTA 3900
 AAGTAAAAAT TAAAAAAAAG TGA

PEZ6 Protein sequence:

Protein Accession #: none

SEQ ID NO:164 PEZ6 DNA SEQUENCE

Nucleic Acid Accession #: AB028945

Coding sequence: 1-3765 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 ATGATGATGA ACCTCCCCGG CGGAGGAGCG GCCGCGGTGA TGATGACGGG CTACAATAAT 60
 GGTCGCTGTC CCCGGAATTC TCTCTACAGT GACTGCATTA TTGAGGAGAA GACGGTGGTC 120
 CTGCAGAAAA AAGACAAATGA GGGCTTTGGA TTCTGTCTTC GAGGGGCCAA AGCTGACACA 180
 CCCATTGAAG AATTCACACC AACACCGGCT TCCCCAGCCC TACAGTACCT GGAGTCCGTG 240
 GATGAAGGTG GGGTGGCGTG GCAAGCCGGA CTAAGGACCG GGGACTTCTT GATTGAGGTT 300
 AACAATGAGA ATGTTGTCAA AGTCGGCCAC AGGCAGGTGG TGAACATGAT CCGGCAGGGA 360
 GGGAAATCACC TGGTCCTTAA GGTGGTCACG GTGACCAGGA ATCTGGACCC CGACGACACC 420
 GCCAGGAAGA AAGTCCCCC GCCTCCAAAG CGGGCACCGA CCACAGCCCT CACCCTGCGC 480
 TCCAAGTCCA TGACCTCGGA GCTGGAGGAG CTCTGGGATA AAGATAAACC CGAGGAGATA 540
 GTCCCGGCCT CCAAGCCCTC CCGCGTCTGT GAGAACATGG CTGTGGAACC GAGGGTGGCG 600
 ACCATCAAGC AGCGGCCAG CAGCCGGTGC TTCCCGGCGG GCTCAGACAT GAACTCTGTG 660
 TACGAACGCC AAGGAATCGC CGTGATGACG CCCACTGTTC CTGGGAGCCC AAAAGCCCCG 720
 TTTCTGGGCA TCCCTCGAGG TACGATGCGA AGGCAGAAAT CAATAGACAG CAGAATCTTT 780
 CTATCAGGAA TAACAGAGGA AGAGCGGCAG TTTCTGGCTC TCCTCAATGCT GAAGTTCACC 840
 AGAAGCCTGT CCAATGCCGA CACCTCTGAG GACATCCCCC CTCCACCGCA GTCTGTGCCC 900
 CCGTCCCCAC CACCACCTTC CCAACCACT TACAACCTGCC CCAAGTCCCC AACTCCAAGA 960
 GTCTACGGGA CGATTAAGCC TGCCTCAAT CAGAATCTG CCGCAAGGT GTCCCCGCC 1020
 ACCAGTCCG ACACCGTGGC CACCATGATG AGGAGAGAAG GATGTACTT CAGGAGAGAG 1080
 CTGGACCGCT ACTCCTTGA CTCTGAAGAC CTCTACAGTC GGAATGCCG CCCGCAAGCC 1140
 AACTTCCGCA ACAAGAGAGG CCAGATGCCA GAAAACCCAT ACTCAGAGGT GGGGAAGATC 1200
 GCCAGCAAG CCGTCTACGT CCCCAGGAG CCGCCAGGC GGAAGGGAT GCTGGTGAAG 1260
 CAGTCCAACG TGGAGGACAG CCCCAGAAAG ACGTGCTCCA TCCCTATCCC GACCATCATC 1320
 GTGAAGGAGC CGTCCACCAG CAGCAGCGGC AAGAGCAGCC AGGGCAGCAG CATGGAGATC 1380
 GACCCCGAG CCGCGGAGCC ACCGAGCCAG CTGCGGCCTG ACGAAAGCCT GACCGTCAGC 1440
 AGCCCTTTG CCGCCGCCAT CGCCGGAGCC GTCCGCGACC GTGAGAAGCG GCTGGAAGCC 1500
 AGGAGGAAT CCCCAGCCTT CCTCTCCACA GACCTGGGGG ATGAGGATGT GGGCCTGGGG 1560
 CCACCGCCC CAGGACCGCG GCCCTCCATG TTCCCGGAGG AGGGGGATT TGTGACGAG 1620
 GACAGCGCTG AGCAGCTGTC ATCCCCCATG CCGAGTGCCA CGCCAGGGA GCCCGAAAAC 1680
 CATTTCTGG GTGGCGCCGA GGCCAGTGCT CCGGGTGAGG CTGGGAGGCC GCTGAATICC 1740
 ACGTCCAAAG CCAAGGGGCC CGAGAGCAGC CCAGCAGTGC CCTCCGCGAG CAGCGGCACA 1800
 GCCGGCCCCG GGAATTATGT CCACCCACTC ACAGGGCGGC TGCTTGATCC CAGTCCCCG 1860
 CTGGCCCTGG CACTGTGCG CAGGACCGA GCCATGAAGG AGTCTCAACA GGGACCCAAA 1920
 GGGGAGGCC CCAAGGCCGA CCTCAACAAA CCTCTTTACA TTGATACCAA AATGCGGCC 1980
 AGCTTGGAT CCGGCTTCCC TACGTACAC AGGCAGAAAC CCGGGGACC CCTGAGGCGG 2040
 CAGGAGACGG AGAACAACTA CGAGACCGAC CTGGGCCGAG ACCGGAAGG CGATGACAAG 2100
 AAGAACATGC TGATCGACAT CATGGACAG TCCAGCAGA AGTCGGGTGG CTTGCTGATG 2160
 GTGCACACCG TGGACGCCAC TAAGTGGAC AACGCCCTGC AGGAAGAGGA CGAGAAGGCA 2220
 GAGGTGGAGA TGAAGGCCA CAGCTCGCCG TCCGAGGTGC CAGAAGGTGT TTCCGAAACC 2280
 GAAGGTGCTT TACAGATCTC CGTGCCCCC GAGCCACCA CCGTCCCCCG CAGAACCATC 2340
 GTCGCGGTGG GCTCCATGGA AGAGGCGGTG ATTTTGCCAT TCCGCATCCC TCCTCCCCCT 2400
 CTGGCATCCG TGGACTTGA TGAGGATTTT ATTTTACAG AGCCATTGCC TCCTCCCCCTG 2460
 GAATTTGCAA ATAGTTTGA TATCCCCGAT GACCGGGCAG CTCTGTCTCC GGCTCTCTCA 2520
 GACTTAGTGA AGCAGAAGAA AAGCGACACC CCTCAGTCCC CTTCGTGAA CTCCAGCCAA 2580
 CCAACCAACT CTGCAGACAG CAAGAAGCCA GCCAGTCTTT CAAACTGTCT GCCTGCCTCA 2640
 TTCTGCCAC CCCCTGAAAG CTTTGACGCC GTCGCGGACT CTGGGATCGA GGAGGTGGAC 2700
 AGCCGGAGTA GCAAGGCCA CCACCTCGAG ACGACCAGCA CTATCTCCAC CGTGTCTAGC 2760
 ATCTCCACCC TGCTTCCGA AGGTGGAGAG AATGTGGACA CCTGCACAGT CTATGCAGAT 2820
 GGGCAAGCAT TTATGTTGA CAAACCCCA GTACCTCTTA AGCCAAAAAT GAAGCCCATC 2880
 ATTCAAAAA GCAATGCACT TTATCAAGAC GCGCTCGTGG AAGAAGATGT AGATAGCTTT 2940
 GTTATCCCC CCGCCGCTCC CCGCCCCCG CCGGGCAGTG CCCAGCCTGG GATGGCCAAG 3000
 GTTCTCAGC CAAGGACCTC CAAGTTGTG GCGACGTCA CAGAGATCAA AAGCCCGATT 3060
 CTCTCAGGCC CAAAGGCAAA CGTTATTAGT GAATTGAACT CTATCCTACA GCAATGAAC 3120
 CGAGAGAAAT TGGCAAGGCC GGGGGAAGGA CTGGATTAC CAATGGGAGC CAAGTCCGCC 3180
 AGCTCGCTC CAAGAAGCCC GGAGATCATG AGCACCATCT CAGGTACACG GAGCACGACG 3240
 GTCACCTTCA CTGTTGCCG CGGCACCTCC CAGCCCATCA CCTGCACAG CCGGCCCCCC 3300
 GACTATGAAA GTCAGACCTC AGGAACAAGA CGTGCCCAA GCCCTGTGGT CTCGCCAACA 3360
 GAGATGAACA AAGAGACCCT GCCCGCCCC CTGTCTGCTG CCACCGCCTC TCCTTCTCCC 3420
 GCTCTCTAG ATGCTTTAG CTTTCCAAGC CAGCCCCCT CTGGGGATCT ATTTGGCTTG 3480
 AACCCAGCGG GACGCACTAG GTCGCCATCC CCTCGATAC TGCAACAGCC AATCTCAAAT 3540
 AAGCCTTTTA CACTAAACC TGTCCACCTG TGGACTAAAC CAGATGTGGC CGATTGGCTG 3600
 GAAAGTCTAA ACTTGGGTGA ACATAAGAG GCCTTCATGG ACAATGAGAT CGATGGCAGT 3660
 CACTTACCAA ACCTGCAGAA GGAGGACCTC ATCGATCTTG GGGTAACCTG AGTCGGGCAC 3720

AGAATGAACA TAGAAAGGGC TTTGAAACAG CTGCTGGACA GATAAGGACG GCTGCTCTCC 3780
 ACCTCGCAGA CTGCTCTTGT TATAAGTAGA GATGGGCTCG TGCTGAAACA TCTGAATGCC 3840
 AAGCGAAGTC TGTGAGCATC AACCCCACTC CATGGGTTTG TCTCTGGTA CCCAAAGAAA 3900
 5 TACTGAGTTG TGTCCACAAC ATGGCTGGGT CTCAGACCC CTGGCTCACC ATGTGGGTGT 3960
 CTTGGGCAGT TTCTATCACA CATGGGACAA GGGGAGGGAG TTTTCTAAC ATGGAAAAAG 4020
 ATTCCAGCC TCCGCCCCAG CATGCAAGTG GCCTCGCTTT GCCGGGTCCG AGAGGCTCCC 4080
 CGTCAATTTT GCACGGGATC CTAGTCTTTG TAGGCAGACA CCAGTGCAC TTAGATACCT 4140
 CCTGAGACCT CCGTCTCTG CTTTCCGGGC AGCTCTCACC ACCCCAGGCC CCGGCATGAG 4200
 10 GCCTTTCCTC AGTCTGTGG CCTCTCAGAG GACACCTGAT GCTCACCTGC CCCTCTTTCT 4260
 CCTGCACTTG GCTTGCAGTG AGATGCTCCC AGATGCATTT GTCCAGTGCC CCATCATGGG 4320
 CCTGAAAGGC AGAGAAACTT TTCTCTACAC AGATCTTTT CCCCATCTCC TCTGTGGTT 4380
 TGCATCCATG GCTCTTTGGC CATGAGGTTT CTGGCAGTGC TGGGAGTTTG GATGGGATCG 4440
 TGCCAGCTT TGCTTAGCTT TCTTTATTTT TGCAAACTCT TTAGCATAAT TCCAAGGTGG 4500
 15 CCAAGCAGAT GTCACATGGA GTTAGTCAAA GCACAAAGTC ACGATTCCAC AATGGAGGGG 4560
 AGACCTGGCC AAGGGAGGCC GCCAGCGTGC AACTGCCCAA GCTCCAGGTC TCCAGGACAA 4620
 GAGCAGTTGT TGTCCATGAG CACCCATCCA GGATGGAGAA TAAGGGCTTC TCTGCCTCTC 4680
 AGAATCTTT TTAATTGAAG ATGTCTTGA GCTTGCAAAG ATCAGAGCAG GTGAGCATCC 4740
 ACTTTGACAT GAAGGACAAG AAGACGCATG GCTCATGGCG GGCACATGCG GCTGCCAGTG 4800
 20 AGACAGCGTC TCTCTGGGA GCTGGGCGGG CACAGCATCC TCAGTTCTGT GCCCAGCCAA 4860
 GGGTGAGCAT CTTGTCTGAG ACAGTCTTT TGCTCTCGGA GGCCAGGGAA GATGGTACTT 4920
 AGAGGCTTTT CCCTATCGC TCTGGGTGTC TAGGAATCCC ACCAGCTTGT CTAAACAGTA 4980
 CAACAGCTTC TTTGAGGACC CAGTGGGTAT GGAGTATAGA CAGAAGCCAG GGTGAGAAC 5040
 AGAAGGTGGG CGGCAGGATC AGAGTGAAG CAGAGGCGTG AGGAGAGGAA AGCAGGGAGG 5100
 25 TCTCTGGGC TGCCAGGTCA GCCTCTCTGG CAAGGCTTTC TTGAGCCCCC CCCTTTTCTT 5160
 TCCCGGAGT CCTCCACCC CATAACAATA CCTCGAATT CCAAAAGAGG TCACCAGATG 5220
 CACATGGGCC GCAAAACACA CAGTCAGGCT TCCAGCACAT TCTCCCCAT TTGGAGGATA 5280
 CTCGAATGTC AGGTTTITGG TTTTATTATT ATTTCAGAAC TAGCTCAGCC CATCTCTAAT 5340
 TATAAAACAT GGTTTTGT TTCTTTTCT CTTTCTTCT TGATTAGGTC TGAACAGCT 5400
 30 CTAGAATGAA CACATAAAAT TTAGCAATTT AAAATCTTTC TTACTGCAA GTTTAAATAG 5460
 TTGTACAGAT AGTTTATAAG CACAATATTT TAAGAAAAAA AAGTGGCTGG TCTACTAGGC 5520
 AGCCTTTGTG CCACCTCAGT CTAGAAAGT TAAAGAAAAA AAAACTTTTG TGATTTAATA 5580
 ATACTATTT TGTGGAATAA TTATAAAAGT ATGACCTTTT TAAATCAACC TTATTTGGAT 5640
 GCATCTGAAC CAGCAGAGCT GTGTTATATT TTCTATCTTT GCTAGAACTT CGTCATTGAA 5700
 35 GGACAATTTT TCAAAAGTGG TTACAATTCA TAATGCAGCA GTTCTTCAA AAACAAAAAC 5760
 AAAACACACA CCACACACAC GCGCTTTTCC AGTCACACAC CCCTGATGTT GGAACCAAGT 5820
 TTTTGGACCT TCTGTTCCAA AACCTTTTGC AGGTCAATCT TTGATTTTGA AATGATCCAA 5880
 TCCAACTTGA AGTCAATTGA ATATTAAGGC GCTTACTTTC CGTGTGCTTT CAGTTTTTCC 5940
 ATCATGAGAT GAATGAGCAT TACTCTAGAT AAATTTCAAG ACAGGATACT ACAGGTGGCC 6000
 40 TGCTGAGGCT GCCCATATTT TTAGAAAAATG TAAAAATGGT GGTITGGCCA TTAATTTGTC 6060
 TTCCATTGTA TGATACCGCA AAATTCCTGT AGTCCATTCC TTTGGCATGG CACTTTCCCT 6120
 GGGCTACAG TTGGTATTAC CTCTGTGCTC AGTGCCAGGC AAAACACTAG CTCAAAGGAG 6180
 AGTCAAGGAA ACCGCTGGCA GACGATAACC AGTCGAAACT CGTGACTTCG GTTTGTTGAA 6240
 CTTTGGCAGC CAGTTGTGTA GGGCCAGATG TTATTCCTT TCTTAAAGAT ACTCCAAGCC 6300
 45 ACATGCCACT AACCACAAGC AAGCTGGCTG CAAGACTAAA GAGCTGATAA CATAGTTTAT 6360
 TTTTACACTG TCTTATTATA GAGAAGTAAT AGACCTATCA GAACCTGCAC TGACCAACAA 6420
 ATAAACACAT GTTGCCAAGA TGAATCGGTC TCTATCTCTA TCTGCTTATT TTGGTACTGA 6480
 AAGCAATAGT TCTCATTTCA AATCACCACC CACTGTCTC CCCTTTGGG ACATGTTAGG 6540
 ACGAGGCCCT ATTCATGCC CCTCTTTAAT GGTGGAACAA ATGTTAAACT GCTCATCTAA 6600
 50 AGATCATGTT GATATTATTC CAGGTTTTAA GATCAACTTT TGTACATAC TGTAATTTAA 6660
 ATAACTGCA TTTACATGCC TAGTTCTGT AATATTGTGT ATACAAAAACC CAAATCTCTC 6720
 AAAATGTAAA TTATGTATAC CTGCCAAGAT ACCTTTTCCA GGGTGTCTGC GCACATTTTA 6780
 AGTTAATTTA CATAATATA AAATTAATCA ATGTGACTGT TGATTGTCTG AACTTTACAT 6840
 ATCACAAGT GAATTATTG TGATACTTTA GTTAATAAAA TGGTAAATTT TTTTCTCAGT 6900
 55 TATTGAACAA GCAAGCATT TCCAGTTGAT TTGGCAATGA CTTTTGTGT GTGGGCCACA 6960
 ATATTGATT TCCCATGAT AATTTTTTTT TGTTTTTTAA ATACTAATAT GTTTCACACT 7020
 ATAGTTTGTG TAACAACACG TGTTCGCATT ATCTATGTTG CTGTTACTTT TGTGCTTTTA 7080
 TTCTTTTGA ACTTTATAAA AAAAAAAAAA AGCTCCTGTA ATTGCACTT TCTCCCAATC 7140
 CTTAAATCTC TTGTATGGCA ACCAAAAATTA CTGTAAAAAA ATAAATATAC TATTGCACTA 7200
 60 AGGTTGTGGT TCTGATTGCA AACAAACAGT GAACACTGTC TGAATTAAAC AAAAAGCTGC 7260
 CCGACTTGCA ATCTAATGTA GATTATCTCA GGCATTGTGG CAGCTCTGTC CTCTCTAAAA 7320
 CTGACCAGAA AAATCTCTCT CATCGAGTAA ACAGGCTCCT GTCAGTGAGC TAATCTGCCT 7380
 TGGTTCCATT TCCTTATTCT CAATTTATCA ATGGATACGT GCATGTTATT TCAGAATTAT 7440
 GCAAAACGTC AAAATCTGCT TCTGTGACCG CTGCTATAGG CGTGGAGCTG AGGCTCGGCT 7500
 65 TTTCTTTTGT TCTGGGTGG AAGCAGCGGT GCCGCGGAGG GCCAGCCAGA TCCGGACCCT 7560
 TCCCTTAGGG TCCAGTCTCC CCACACCCA GCAGGGTGTCT TTCTAGCCAT AAGGCCAAGG 7620
 GAGTGGCAGA ACTGGGCCGC CTCTCTGGTT GACAAGCAAA CCACATGCTA AGGCTTGGAG 7680
 CAAGAGAGAA TTTGTGTCTA TTGGCAAAGA ACTAAGCCAG GAAGACATGG GCCATCCCTC 7740
 CGCTTAGGG AAGCAATTTT TAAACCTAAA CGTGAACCTT CTCTTTGGC CTCACCAAGT 7800
 70 AAAACTTGTG GTCTTAGTT CCTAAAGTTT CTCTACTTT GGCACATTCC CCAGTTGAGC 7860
 AGCAGCTCT ATGCTTCCAC GTTCAGGAAA AATTCAGTC CTCATATCTT TTGTAGTTCA 7920
 CCTCAAGCT TCCCGCTTC ACCATCCAAAT AGTTTCTCCC AAACCTTGGC ACCCCCTAG 7980
 ACTTGTCTC CAATGGTTTC TTCCAGACCA CTTTCTAG ATGAATATAT TCGTTTACCT 8040
 TACTAGGAAA ATTTATGGAA GATTTTTCT TTTACTTGA ATTGGAGCA TTITAATAAC 8100
 75 TGGCGAAGT GAATGTGTTT CTGTATTTGT AGACAACCAT GTACCCATGC AAGTAGGTGA 8160
 ACATTCCACA GTGGCTGGGT GACCACAGCA GCTGCATGCA GACAGGACTG CCCGTGCTTT 8220
 GTGGGAATC AGAGAAATTC CAAACTTGT TCTCAGACTT CCGCAGATCT CATCACTTTG 8280
 ATTCTAATC CATGCTGAT TGGTGATTTT GTTATCTGTT CCTGTAACCT GTTCTACATT 8340
 CCACAGTCTT TACCGTTTAA TGTTCAAAAT TACAACAATC CCTGTCCATT GATTCCACTC 8400
 TGGAACCTCT TGTTCATGCC AATTTTGAAA TTTAATACG AGCCTTCAA TAAACACAGA 8460

AAAGAAAAAA AAAAAAAAAA AAAAAAAAAA

SEQ ID NO:165 PEZ6 Protein sequence;
Protein Accession #: BAA82974.1

1 11 21 31 41 51
| | | | |
5 10 15 20 25 30
MMMNVPVGGGA AAVMMTGYYN GRCPRNSLYS DCIIEEKTVV LQKKNDEGFG FVLRGAKADT 60
PIEFTPTPA FPALQYLESV DEGGVAWQAG LRTGDFLIEV NNENVVKVGH RQVVNMIRQG 120
GNHLVLKVVT VTRNLDPDDT ARKKAPPPPK RAPTTALT LR SKSMTSELEE LVDDKDKPEEI 180
VPASKPSRAA ENMAVEPRVA TIKQRPSSRC FPAQSDMNSV YERQGI VMT FTVPGSPKAP 240
FLGIPRGTM RQKSIDSRI FLSGITEERQ FLAPPM LKFT RLSMPD TSE DIPPPQSV P 300
PSPPPSPPT YNCPKSPPT R VYGTIKPAFN QNSAAKVSPA TRSDTVATMM REKGM YFRRE 360
LDRLSLDSE LYSRNAGPQA NFRNKRQGM P ENPYSEVGKI ASKAVYVPAK PARRKGMLVK 420
QSNVEDSPEK TCSIPIPTI VKEPSTSSG KSSQGSSMEI DPQAPPEPSQ LRPDES LTVS 480
SPFAAAIAGA VRDREKRLEA RRNSPAFLST DLGDEDVGLG PPAPRTRPSM FPEGDFADE 540
DSAEQLSSPM PSATPREPEN HFVGGAEASA PGEAGRPLNS TSKAQGPSS PAVPSASSGT 600
AGPGNYVHPL TGRLLDPSSP LALALSARDR AMKESQGGPK GEAPKADLNK PLYIDTKMRP 660
SLDAGFTVT RQNTGRLRR QETENKYETD LGRDRKGDDK KNMLIDIMDT SQQKSAGLLM 720
VHTVDATKLD NALQEEDEKA EVEMKPDSSP SEVPEGVSET EGALQISAAP EPTTVPGRTI 780
VAVGSMEEAV ILPFRIPPPP LASVDLDEDF IFTEPLPPPL EFANSFDIPD DRAASVPALS 840
DLVKQKKS DT PQSPSLNSSQ PTNSADSKKP ASLSNCLPAS FLPPPE SFD VADSGIEEVD 900
25 SRSSSDHILE TTSTISTVSS ISLSSEGGE NVDICTVYAD GQAFMVDKPP VPPKPKMKPI 960
IHKSNALYQD ALVEEDVDSF VIPPPAPPPP PGSAQPGMAK VLQPRTSKLW GDVTEIKSPI 1020
LSGPKANVIS ELNSILQOMN REKLAKPGE G LDSPMGAKSA SLAPRSPEIM STISGTRSTT 1080
VTFTVRPGTS QPITLQSRPP DYESTSGTR RAPSPVVSPT EMNKETLPAP LSAATASPS P 1140
30 ALSDVFSLPS QPPSGDLFGL NPAGRSRSPS PSILQQPISN KPFTTKPVHL WTKPDVADWL 1200
ESLNLGEHKE AFMDNEIDGS HLPNLQKEDL IDLGVTRVGH RMNIERALKQ LLD R

SEQ ID NO:166 PEZ4 DNA SEQUENCE

Nucleic Acid Accession #: NM_000024
Coding sequence: 220-1461 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
| | | | |
35 40 45 50 55 60 65 70
ACTGCGAAGC GGCTTCTTCA GAGCACGGGC TGGAAC TGGC AGGCACCGCG AGCCCTAGC 60
ACCCGACAAG CTGAGTGTGC AGGACGAGTC CCCACCACAC CCACACCACA GCCGCTGAAT 120
GAGGCTTCCA GCGTCCGCT CGCGGCCCGC AGAGCCCCGC CGTGGGTCCG CCCGCTGAGG 180
CGCCCCCAGC CAGTGCCTT ACCTGCCAGA CTGCGCGCCA TGGGGCAACC CGGGAACGGC 240
AGCGCCTTCT TCGTGGCACC CAATAGAAGC CATGCGCCGG ACCACGACGT CACGCAGCAA 300
AGGGACGAGG TGTGGTGGT GGCATGGGC ATCGTCATGT CTCTCATCGT CCTGGCCATC 360
45 GTGTTTGGCA ATGTGCTGGT CATCACAGCC ATTGCCAAGT TCGAGCGTCT GCAGACGGTC 420
ACCAACTACT TCATCACTTC ACTGGCCTGT GCTGATCTGG TCATGGGCCT GGCAGTGGTG 480
CCCTTTGGGG CCGCCCATAT TCTATGAAA ATGTGGACTT TTGGCAACTT CTGGTGCGAG 540
TTTGGGACTT CCATTGATGT GCTGTGCGTC ACGGCCAGCA TTGAGACCTT GTGCGTGATC 600
CGAGTGGATC GCTACTTGC CATTACTTCA CCTTTCAAGT ACCAGAGCCT GCTGACCAAG 660
50 AATAAGGCC CCGGATCAT TCTGATGGTG TGGATTGTGT CAGGCCTTAC CTCCTTCTTG 720
CCCATTGAGA TGCACTGGTA CCGGGCCACC CACCAGGAAG CCATCAACTG CTATGCCAAT 780
GAGACCTGCT GTGACTTCT CACGAACCAA GCCTATGCCA TTGCCTCTTC CATCGTGTCC 840
TTCTACGTTT CCCTGGTGAT CATGGTCTTC GTCTACTCCA GGGTCTTTCA GGAGGCCAAA 900
AGGCAGCTCC AGAAGATTGA CAAATCTGAG GGCCGCTTCC ATGTCCAGAA CCTTAGCCAG 960
55 GTGGAGCAGG ATGGGCGGAC GGGGCATGGA CTCCGCAGAT CTTCCAAGTT CTGCTTGAAG 1020
GAGCACAAG CCCTCAAGAC GTTAGGCATC ATCATGGGCA CTTTACCCT CTGCTGGCTG 1080
CCCTTCTTCA TCGTTAATAT TGTGCATGTG ATCCAGGATA ACCTCATCCG TAAGGAAGTT 1140
TACATCCTCC TAAATGGAT AGGCTATGTC AATTCTGGTT TCAATCCCT TATCTACTGC 1200
CGGAGCCAG ATTTAGGAT TGCCTTCCAG GAGCTTCTGT GCCTGCGCAG GTCTTCTTTG 1260
60 AAGCCTATG GGAATGGCTA CTCCAGCAAC GGCAACACAG GGGAGCAGAG TGGATATCAC 1320
GTGGAACAGG AGAAAGAAAA TAACTGCTG TGTGAAGACC TCCAGGCAC GGAAGACTTT 1380
GTGGGCCATC AAGTACTGT GCCTAGCGAT AACATTGATT CACAAGGGAG GAATTGTAGT 1440
ACAAATGACT CACTGCTGTA AAGCAGTTTT TCTACTTTTA AAGACCCCCC CCCCCCAAC 1500
65 AGAACACTAA ACAGACTATT TAACTGAGG GTAATAAACT TAGAATAAAA TTGTAATAAT 1560
TGTATAGAGA TATGCAGAAG GAAGGGCATC CTCTGCGCTT TTTTATTTTT TTAAGCTGTA 1620
AAAAGAGAGA AAACCTATT T GAGTGATTAT TTGTTATTG TACAGTTTCA GTTCTCTTTG 1680
CATGGAATTT GTAAGTTTAT GTCTAAAGAG CTTTAGTCTT AGAGGACCTG AGTCTGCTAT 1740
ATTTTCATGA CTTTTCATAT TATCTACCTC ACTATTCAAG TATTAGGGGT AATATATTGC 1800
70 TGCTGGTAAT TTGATCTGA AGGAGATTTT CCTTCTACA CCCTTGGACT TGAGGATTTT 1860
GAGTATCTCG GACCTTTCAG CTGTGAACAT GGACTCTTCC CCCACTCTC TTATTGCTC 1920
ACACGGGGTA TTTTAGGCA GGAATTGAGG AGCAGCTTCA GTTGTTTTC CGAGCAAAGG 1980
TCTAAAGTTT ACAGTAAATA AAATGTTTGA CCATG

SEQ ID NO:167 PEZ4 Protein sequence;
Protein Accession #: NP_000015.1

1 11 21 31 41 51

| | | | |
 MGQPGNGSAF LLAPNRSHAP DHDVTQQRDE VVVVGMGIVM SLIVLAIVFG NVLVITAIK 60
 FERLQTVTNY FITSLACADL VMGLAVVPFG AAHILMKMWT FGNFWCEFWT SIDVLCVTAS 120
 IETLCVIAVD RYFAITSPFK YQSLLTKNKA RVILMVWIV SGLTSFLPIQ MHWYRATHQE 180
 AINCYANETC CDFFTNQAYA IASSIVSFYV PLVIMVFVYS RVFQEAQRQL QKIDKSEGRF 240
 HVQNLSQVEQ DGRGTGHGLRR SSKFCLKEHK ALKTLGIIMG TFTLCWLPFF IVNIVHVIQD 300
 NLRKEVYL LNWIGYVNSG FNPLIYCRSP DFRIAFQELL CLRRSSLKAY GNGYSSNGNT 360
 GEQSGYHVEQ EKENKLLCED LPGTEDFVGH QGTVPDNDID SQGRNCSTND SLL

SEQ ID NO:168 PEZ1 DNA SEQUENCE

Nucleic Acid Accession #: NM_004457
 Coding sequence: 143-2305 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 GAATTCGTTG TTGGGAAGGA CTGGGGAAC AGCTGTAACA TTGCCACCC TCAGAAGCTG 60
 CTGGTCCTGT GTCACACCAC CTTAGCCTCT TGATCGAGGA AGATTCTCGC TGAAGTCTGT 120
 TAATTCTACT TTTTGAGTAC TTATGAATAA CCACGTGTCT' TCAAAACCAT CTACCATGAA 180
 GCTAAACAT ACCATCAACC CTATTCTTTT ATATTTTATA CATTITCTAA TATCACTTTA 240
 TACTATTTTA ACATACATTC CGTTTTATTT TTCTCCGAG TCAAGACAAG AAAAATCAAA 300
 CCGAATTAAG GCAAAGCCTG TAAATTCAAA ACCTGATTCT GCATACAGAT CTGTTAATAG 360
 TTTGGATGGT TTGGCTTCAG TATTATACCC TGGATGTGAT ACTTTAGATA AAGTTTTTAC 420
 ATATGCAAAA AACAAATTTA AGAACAAAAG ACTCTTGGGA ACACGTGAAG TTTTAAATGA 480
 GGAAGATGAA GTACAAACCA ATGGA AAAAT TTTTAAAAAG GTTATTCTTG GACAGTATAA 540
 TTGGCTTTC TATGAAGATG TCCTTGTCG AGCCTTTAAT TTTGAAAATG GATTACAGAT 600
 GTTGGGTCAG AAACCAAAGA CCAACATCGC CATCTTCTGT GAGACCAGGG CCGAGTGGAT 660
 GATAGCTGCA CAGGCGTGT TTATGTATAA TTTTCAGCTT GTTACATTAT ATGCCACTCT 720
 AGGAGGTCCA GCCATTGTTT ATGCATTAAG TGAACACAGG GTGACCAACA TCATTACTAG 780
 TAAAGAACTC TTACAAACAA AGTTGAAGGA TATAGTTTCT TTGGTCCCAC GCCTGCGGCA 840
 CATCATCACT GTTGATGGAA AGCCACCGAC CTGGTCCGAC TTCCCAAGG GCATCATTGT 900
 GCATACCATG GCTGCAGTGG AGGCCCTGGG AGCCAAGGCC AGCATGGAAA ACCAACCTCA 960
 TAGCAAACCA TTGCCCTCAG ATATTGCAGT AATCATGTAC ACAAGTGGAT CCACAGGACT 1020
 TCCAAAGGGA GTCATGATCT CACATAGTAA CATTATTGCT GGTATAACTG GGATGGCAGA 1080
 AAGGATTCCA GAACTAGGAG AGGAAGATGT CTACATTGGA TATTTGCTC TGCCCATGT 1140
 TCTAGAATTA AGTGTGAGC TTGTCTGTCT TTCTCACGGA TGCCGCAATT GTTACTTCTC 1200
 ACCACAGACT TTAGCAGATC AGTCTTCAAA AATTA AAAAAA GGAAGCAAAG GGGATACATC 1260
 CATGTTGAAA CCAACACTGA TGGCAGCAGT TCCGGAATC ATGGATCGGA TCTACAAAAA 1320
 TGTCATGAAT AAAGTCAGTG AAATGAGTAG TTTTCAACGT AATCTGTTTA TTCTGGCCTA 1380
 TAATTACAAA ATGGAACAGA TTCAA AAGG ACGTAATACT CCACTGTGCG ACAGCTTTGT 1440
 TTTCGGGAAA GTTCGAAGCT TGCTAGGGGG AAATATTCTG CTCTGTTTGT GTGGTGGCGC 1500
 TCCACTTTCT GCAACCACGC AGCGATTCT GAACATCTGT TTCTGCTGTC CTGTTGGTCA 1560
 GGGATACGGG CTCACTGAAT CTGCTGGGGC TGAACAATT TCCGAAGTGT GGGACTACAA 1620
 TACTGGCAGA GTGGGAGCAC CATTAGTTTG CTGTGAAATC AAATTA AAAAAA ACTGGGAGGA 1680
 AGGTGGATAC TTTAATACTG ATAAGCCACA CCCCAGGGGT GAAATCTTAA TTGGGGGGCCA 1740
 AAGTGTGACA ATGGGGTACT ACAAAAATGA AGCAAAAACA AAAGCTGATT TCTCTGAAGA 1800
 TGAAAATGGA CAAAGGTGGC TCTGTACTGG GGATATTGGA GAGTTTGAAC CCGATGGATG 1860
 CTTAAAGATT ATTGATCGTA AAAAGGACCT TGTA AAAAAA CAGGCAGGGG AATATGTTTC 1920
 TCTTGGGAAA GTACAGGCAG CTTTGAAGAA TCTTCCACTA GTAGATAACA TTTGTGCATA 1980
 TGCAACAGT TATCATTTCT ATGTCATTGG ATTGTGTTGT CCAATCAAA AGGAACCTAAC 2040
 TGAAGTAGCT CGAAGAAAG GACTTAAAGG GACTTGGGAG GAGCTGTGTA ACAGTTGTGA 2100
 AATGGAAAAA GAGGTACTTC AAGTGCTTTC CGAAGCTGCT ATTTAGCAAA GTCTGGAAAA 2160
 GTTTGAAATT CCAGTAAAAA TTCGTTGAG TCCTGAACCG TGGACCCCTG AAAGTGGTCT 2220
 GGTGACAGT GCCTTCAAGC TGAACGCAAG AGAGCTTAAA ACACATTACC AGGCGGACAT 2280
 TGAGCGAATG TATGGAAGAA AATAATTATT CTCTCTGGC ATCAGTTTGC TACAGTGAGC 2340
 TCACATCAAA TAGGAAAAATA CTGAAATGC ATGTCTCAAG CTGCAAGGCA AACTCCATTCT 2400
 CTCATATTA ACTATTACTT CTCATGACGT CACCATTTTT AACTGACAGG ATTAGTAAAA 2460
 CATTAGACA GCAAACCTGT GTCTGTCTCT TCTTTCATTT TCCCGCCAC CAACTTACTT 2520
 TACCACCTAT GACTGTACTT GTCAGTATGA GAATTTTCT GAATCATATT GGGGAAGCAG 2580
 TGATTTTAAA ACCTCAAGTT TTTAAACATG ATTTATATGT TCTGTATAAT GTTCAGTTTG 2640
 TAACITTTTA AAAGTTTGA TGTATAGAGG GATAAATAGG AAATATAAGA ATTGGTTATT 2700
 TGGGGGCTTT TTACTTACT GTATTTAAAA ATACAAGGGT ATTGATATGA AATTATGTAA 2760
 ATTCAAATG CTATGAATC AAATCATTGT TGAACAAAAG ATTGTGTGCT GTGAATTTAT 2820
 TGTCTGTAT GCATTTGAGA GAAATAAATA TACCCATACT TATGTTTAA GAAGTTGAGA 2880
 TCTGTGAAT ATATGCTGT CAGTGTCTTC TTTATATATT TATTTTAT TAGAAAAAAT 2940
 GAAGTTTGGT TGGTGATGCA TGAACAAAA TAGCAAGAGA GGGTTATAGT TTAATAGTAA 3000
 GGGAGATAAC ACAGCATGTG TAGCACCAGT TGATAATTGG TCTCTAGTAG CTACTGTCA 3060
 AAATGTTCAA TGAAGTCTTC TGTTCATCTG TTGAACTAG GAAAAATACC AAACITAAAT 3120
 GGAAGAAATC TGAAGAGAG GATAGAATTT AAAGAACAAG AGTATATAAA GTTATTCTTT 3180
 GAATATTTCG TTGACTATT GTACATTGAG TTATCTATAT TTGTAAACAA ATTAGTCATG 3240
 GAAAATTATT CTATTCCAAA GTCTCCTTT AGTCTAGATA ATCATTTT CATTTTAAAA 3300
 TTAGTGTITT TCATAGTTTG CACTGATGCG TGTATGGATG TGTGTGAGTC AGTGGTAGCT 3360
 TATTTAAAAA GCACCTTCTC TTTCTCCA TAACCTTTGT AACTAAAAAA ATGAAAGAAT 3420
 TTAGAATGTA TTTGATGATA GCATTCTCAC TAAGACACAT GAGAATTTAA CTTTATAACC 3480
 GCGTGAGTTA AGATTTAATT CATAGGTTTT GATGTCATTG TTGAAGTTAT TTGTAATCA 3540
 GAAACCTTGC TTGTGTGATA CATAGTAAGT CTCTTCATT ATTACTGCTT GCCTGTGTTT 3600

ATATCTGGAT TATCAAAAGC AATAGTGCAC CAATTAAGAT GTGCTCAAAT CAGGACTTAA 3660
 ATCATAGGCA CCACATTTTT CATGTCAGAC TAGTTACTTT GTTGATTCTC AGTTACTGTA 3720
 GGCATCAAAA GGCAAAATC A

SEQ ID NO:169 PEZ1 Protein sequence:
 Protein Accession #: NP_004448.1

1 11 21 31 41 51
 MNHVSSKPS TMKLKHTINP ILLYFIHFLI SLYTILTYIP FYFFSESROE KSNRIKAKPV 60
 NSKPD SAYRS VNSLDGLASV LYPGCDTLDK VFTYAKNKF NKRLLTREV LNEDEVQPN 120
 GKIFKKVILG QYNWLSYEDV FVRAFNFGNG LQMLGQPKPT NIAIFCEIRA EWMIAAQACF 180
 MYNFQLVTLY ATLGGAIVH ALNETEVTNI ITSHELLQTK LKDIVSLVPR LRHIITVDGK 240
 PPTWSDFPKG IIVHTMAAVE ALGAKASMEN QPHSKPLPSD IAVIMYTSGS TGLPKGVMIS 300
 HSNIIAGITG MAERIPELGE EDVYIGYLP L AHVLELSAEL VCLSHGCRIG YSSPQTLADQ 360
 SSKIKKSGK DTSMLKPTLM AAVPEIMDRI YKNVMNKVSE MSSQRNLFI LAYNYKMEQI 420
 SKGRNTPLCD SFVFRKVRSL LGGNIRLLC GGAPLSATTQ RFMNICFCCP VGQGYGLTES 480
 AGAGTISEVW DYNVTGRVGAP LVCCEIKLN WEEGGYFNTD KPHPRGEILI GGQSVTMGY 540
 KNEAKTKADF SEDENQQRWL CTGDIGEFEP DGCLKIDRK KDLVKLQAGE YVSLGKVEAA 600
 LKNLPLVDNI CAYANSYHSY VIGFVVPNQK ELTELARKKG LKGTWEELCN SCEMENEVLK 660
 VLSEAAISAS LEKFEIPVKI RLSPEPWTP E TGLVTD AFKL KRKELKTHYQ ADIERMYGRK

SEQ ID NO:170 PCQ7 DNA SEQUENCE
 Nucleic Acid Accession #: none found
 Coding sequence: 38-1075 (underlined sequence corresponds to start and stop codon)

1 11 21 31 41 51
 AGCAACGACG CCGGGCAGCG GGAGCGGCGG CCGCGCCATG TGGCTGCTGG GGCCGCTGTG 60
 CCTGCTGCTG AGCAGCGCGG CCGAGAGCCA GCTGCTCCCC GGGAAACAAT TCACCAATGA 120
 GTGCAACATA CCAGGCAACT TCATGTGCAG CAATGGACGG TGCATCCCGG GCGCCTGGCA 180
 GTGTGACGGG CTGCTGACT GCTTCGACAA GAGTGATGAG AAGGAGTGCC CCAAGGCTAA 240
 GTCGAAATGT GGCCCAACCT TCTTCCCTTG TGCCAGCGGC ATCCATTGCA TCATTGGTCG 300
 CTTCGCGTGC AATGGGTGTG AGGACTGTCC CGATGGCAGC GATGAAGAGA ACTGCACAGC 360
 AAACCCCTCTG CTTTGTCTCCA CCGCCCGCTA CCACTGCAAG AACGGCCTCT GTATTGACAA 420
 GAGCTTCATC TGCGATGGAC AGAATAACTG TCAAGACAAC AGTGATGAGG AAAGCTGTGA 480
 AAGTTCTCAA GAACCCGGCA GTGGCAGGT GTTTGTGACT TCAGAGAACC AACTTGTGTA 540
 TTACCCAGC ATCACCATG CCAATCATCG CAGCTCCGTC ATTTTGTGTC TGGTGGTGGC 600
 CCTGCTGGCA CTGCTCTGTC ACCACGAGCG GAAGCGGAAC AACCTCATGA CGCTGCCCGT 660
 GCACCGGCTG CAGCACCCTG TGCTGCTGTC CCGCCTGGTG GTCCTGGACC ACCCCACCA 720
 CTGCAACGTC ACCTACAACG TCAATAATGG CATCCAGTAT GTGCCAGGCC AGCGGAGCA 780
 GAATGCGTCG GAAGTAGGCT CCCCACCTTC TACTCCGAG GCCTTGTCTG ACCAGAGGCC 840
 TGCCTGGTAT GACCTTCTCT CACCGCCCTA CTCTCTGAC ACGGAATCTC TGAACCAAGC 900
 CGACCTGCCC CCGTACCGCT CCGGCTCCGG GAGTGCCAAAC AGTGCCAGCT CCCAGGCAGC 960
 CAGCAGCCTC CTGAGCGTGG AAGACACCAG CCACAGCCCG GGGCAGCCTG GCCCCAGGA 1020
 GGGCACTGCT GAGCCAGGG ACTCTGAGCC CAGCCAGGGC ACTGAAGAAG TATAAGTCCC 1080
 AGTTATTCCA AAGTCCATAT GGGTTAATCT GCTCTGACTT GTTGCCATTC TAACAATTG 1140
 TGCTCATGGG AAGCTCTTTA AGCACCTGTA AGGATGTCTC AAGTTACAGT TTGGGATATT 1200
 AACTATCTCT GCATTCCCTT CCTCCCCAG ACTTCAGAGA TGTTTCTCTG GCGTCTCAGT 1260
 TGACATGATC TGTGTGTCGT CTTTCTGTG AGGTCACTCT TCCCTTGGGA CCCGAGATCA 1320
 CACCTCATTT TTTTCACATTA TTCTGTTTCT GTTGAGAGA CAGCATATAA AACAGTATTG 1380
 AAATAGGCTG GGAGAGAGCA ATGTTTCTGT GCTATATTGG ATGCTCAGAA GTGCAGGAGA 1440
 CGCTGGACCC AATTCTCTCT GCTGGGTAGT TACCTTATAG CATTTGGGGA TTTGGGTTAG 1500
 ATGATCTAAC CAGGAGGCCA TCACTGGATG GTCACCCCCC CAAAAAATT CCATTGTAGC 1560
 ATCAAAACCT GCTTGCACA ATCTATTTG ATGCCCCCAG TTCAGCAGAG TCAGTGGCCA 1620
 AAGAAAACCT TGGACGTGAG TAACACCCTT CAGCAGTCGC AACGTTATTT TGGTTTGTG 1680
 AAGGACTCTG AAACCATCTA CCTGTATATA ATTCTGGCTT TAGAAATTTG CCAAGAATG 1740
 CTCATTCTGA GAGCTTCTCT CAGCAGCATA TATCATCAGC CTCATCTTAA AATAGGCAGG 1800
 GAGCCCTCC CATGAGTTTA TCCAAGTTCT CAGCTCCTAA AATGCAGGCT GCCAAGACCC 1860
 TACACCTGCC CTGGCTCTAC AGCCACTTAC CTGGTTTCTG GACTGTCACC CTCCCAGCTG 1920
 ACCTGCCCGT AGCCAAGGAA TGAGGACCTA ACTTGAGTTG GCCCAAAGTC TGACCTGGCT 1980
 GTATGTCCTT GTGGCCACA CCCAGCCTGT CTTGCTCATT CATGCAGCCT CAACACTGGC 2040
 CTCCAAAGTT CCCTTAACAC TTGCAAAGTC CTTTTCACCT GTGCATTTGG ACTTGAGGAC 2100
 ACTGTTTCT ATCAGAGTG AGAGCCATGT TCAATACCTC CAGCAAGCTC TCCTGGCTCC 2160
 CTGCACTGTG CAGCTCTCTC TTCCCAAGGT CCCAATACCA GCACCTCTAG TTAGAGTTAG 2220
 GGTCAAGGCT AGGCCTCTCC CAACATCCCA GTAGTTTCTC CTCTGAGACA CATGGGCAAG 2280
 AGACAATTTG GAGTCAAGAT TTTCCATTGT GATCTATTTT AAATCTTTTA GAAATGCATT 2340
 TGAAACAGTG TGTFTGTTTT TTCCCTTCTA GTTAAGGGAC TATTATATATG TGTATAGGAA 2400
 AGCTGTCTCT TTTTGTGTTT TTCTTTAAAC AAGGTCCAAA GAAAGATGCA AAAGGAGATC 2460
 ACACCTTTGC CCGCTGTGAG CCGGTGATAA CAAGTCACTC CAGACTAACC TGTGTGCCAG 2520
 ACATTTGTGC ATTGTTGCAC TTTGAGGTTA TTATTATCA AGTTCTTGAA GGAAGCAGAA 2580
 AGAGGGACTC CTCCTCTCCCT CCGGTATAG TCTCTATGTT TGTGCTAGTT TTTCTTTTTT 2640
 TTCTCTGTGT CCACTGAGCC ACAGGGCCCC CCTCCCTGCA GGAATAAGGG GTAAAACGTT 2700
 AGGTGTTGTT TGGCAAGAAA CCACACTGAC TGATGAGGGG TAAATAGGAA CCAGGTAGAG 2760
 CCACCTCCGG CAGCTGTGAC CCAATCAGAA CTTCTTTCCG CAGCTGAAGA AATGTTTCAGT 2820
 AACCCTGTTT ACCTAATTA AAACAGAGCC TGCAGGAAGT GGGGCTAAAG TGGCATTCAG 2880
 TGATCCTGTT CTCTTCTTTT TTTTAAACCA ATCAAAGGA TGTACAGAA 2940

5	AAGCTAGCCA	CTGGTATTTT	GTTTGTGTTA	AAAAAAAAAA	GAAAGAAAGA	AAGAAAGAAA	3000
	AACGGAAAGG	AACTAGACTG	CCTGTATCTT	TCATTTTAA	AATAGCACTT	GAGTTATTTT	3060
	CTGAGTAATC	CAATAAAGAA	CTTTTGATGA	CAGCCAGAAT	GTGTTAGAAC	TCTGGCTGAA	3120
	CATTTCATCT	CCTGTGAGTC	AGAAGGGCTT	TATTTCTCCC	TTTGTATGGG	CCCTTCTTC	3180
	TTTCTGGTGC	TCTGGAAGTT	GTTTAGAGGA	AAGAATTCCTA	ATTTTAATTA	ATTGCGCAGT	3240
	GAGTTAATCT	CATCTCGCTT	TCTGCTTCCA	GGCATCTTAG	GAAAAACAAA	TGGTTTATAGT	3300
	AGATAAGGGA	TGCTACTATA	TGCTTTTATA	AAACAAACAG	GGACATTTT	ATTATAGATT	3360
	TGATTTT	AATGAATGTT	TTTAAAAATA	TATAAATAGG	ACACCAAAGC	GGCAGGTTT	3420
10	TTTTTGGGGG	GAGGGGGTTT	GTTTCCCAAC	TCAAGATGGC	ACATTAGTGG	CCAGCAATAT	3480
	TTTTTAACTC	ATTCCAACCA	GGAAGCTTTT	TTATACATTG	CCTAAATCTA	CGCCAACCCAG	3540
	AAAATAGTCT	CATCTCTTTT	TTTCTCAAA	GAGATCCGTG	TTTATTTT	GCATTAAATT	3600
	AGTTACACTG	TGATGACTGG	CCTATTACCT	GACTCAGCTC	CCTCTACCTT	GAAATTGACA	3660
	TTTTTAAAAA	ATGCAACTAA	GTGGTTAATA	GTGTGTGACG	CTCAAAGTTA	ATGTAAACTG	3720
15	GAAAGGTTGT	GTGTCGTGTC	TTTTTGTGTT	TTGGTTAGGC	TTGGTTTGT	TTTTTAATTT	3780
	TTATACTTTC	TAATAAATTT	GCAGTTTCAT	TCTTCTGTTT	TGTGCAAAWG	GWMCTAMARM	3840
	AAMMAAAAC	AWYWTGGGG	GGGCTTGGGC	CTCGGAAAAA	GTTTTTAACA	CCACTTCGGG	3900
	TGGGGCGCGG	GGGCCACCGT	AGGTACGCGG	ACCACGCGGG	CCCAAACGGG	ACCCCAAGAG	3960
	GAAACCCCTG	CCAAGAAAAA	GGTGGCGAGA	ATCTCCACAC	CCAGAAAAAA	ACGCGCCGGG	4020
20	GGAAACCGCA	GAGTGTTCGG	TAAACACAC	CCGAAGAGAG	AACCTCAGAAG	CACACAAGCG	4080
	GGACTCAACC	AGGAGGACCC	AAGGGAACCC	GATAGAGTAC	G		

25 SEQ ID NO:171 PCQ7 Protein sequence:
Protein Accession #: none found

30	1	11	21	31	41	51	
	MWLLGPLCLL	LSSAAESQLL	PNNFTNECN	IPGNFMCSNG	RCIPGAWQCD	GLPDCFDKSD	60
	EKECPKAKSK	CGPTFFPCAS	GIHICIIRFR	CNGFEDCPDG	SDEENCTANP	LLCSTARVHC	120
	KNLGLCIKSF	ICDGNQNCQD	NSDEESCESS	QEPGSGQVVF	TSENQLVYYP	SITYALIGSS	180
	VIFVLVUALL	ALVLHHQRKR	NNLMTLPVHR	LQHPVLLSRL	VVLDPHHCN	VTYNVNMGIQ	240
35	YVASQABQNA	SEVGSPPSYS	EALLDQRPFW	YDLPPPPYS	DTESLNQADL	PPYRSRSGSA	300
	NSASSQAASS	LLSVEDTSHS	PGQPGPQEGT	AEPRDSEPSQ	GTEEV		

40 SEQ ID NO:172 PEL3 DNA SEQUENCE
Nucleic Acid Accession #: NM_005856.1
Coding sequence: 57-1535 (underlined sequences correspond to start and stop codons)

45	1	11	21	31	41	51	
	GTCTATTTGA	ACATTCCAGA	TACCTATCAT	TACTCGATGC	TGTTGATAAC	AGCAAGATGG	60
	CTTTGAACCT	AGGGTCACCA	CCAGCTATTG	GACCTTACTA	TGAAAACCAT	GGATACCAAC	120
	CGGAAAACCT	CTATCCCGCA	CAGCCCACTG	TGGTCCCCAC	TGCTTACGAG	GTGCATCCGG	180
	CTCAGTACTA	CCCGTCCCCC	GTGCCCCAGT	ACGCCCCGAG	GGTCTTGACG	CAGGCTTCCA	240
	ACCCCGTCGT	CTGCACGCG	CCCAATCCCT	CATCCGGGAC	AGTGTGCACC	TCAAAGACTA	300
50	AGAAAGCACT	GTGCATCACC	TTGACCCTGG	GGACCTTCCT	CGTGGGAGCT	CGGCTGGCCG	360
	CTGGCCTACT	CTGGAAGTTC	ATGGGCAGCA	AGTGCTCCAA	CTCTGGGATA	GAGTGCAGCT	420
	CCTCAGGTAC	CTGCATCAAC	CCCTCTAACT	GGTGTGATGG	CGTGTACAC	TGCCCCGGCG	480
	GGGAGGACGA	GAATCGGTGT	GTTTCGCTCT	ACGGACCAAA	CTTCATCCCT	CAGATGTACT	540
	CATCTCAGAG	GAACTCCTGG	CACCCCTGTG	GCCAAGACGA	CTGGAACGAG	AACTACGGGC	600
55	GGCGGGCCTG	CAGGGACATG	GGCTATAAGA	ATAATTTT	CTCTAGCCAA	GGAAATAGTGG	660
	ATGACAGCGG	ATCCACCAGC	TTTATGAAAC	TGAACACAAG	TGCCGGCAAT	GTCTGATATCT	720
	ATAAAAAACT	GTACCACAGT	GATGCCTGTT	CTTCAAAAGC	AGTGGTTTCT	TTACGCTGTT	780
	TAGCCTGCGG	GGTCAACTTG	AACTCAAGCC	GCCAGAGCAG	GATCGTGGGC	GGTGAGAGCG	840
	CGCTCCCGGG	GGCCTGGCCC	TGGCAGGTCA	GCCTGCACGT	CCAGAACGTC	CACGTGTGCG	900
60	GAGGCTCCAT	CATCACCCCT	GAGTGGATCG	TGACAGCCGC	CCACTGCGTG	GAAAAACCTC	960
	TTAACAATCC	ATGGCATTTG	ACGGCATTTG	CGGGGATTTT	GAGACAATCT	TTCATGTTCT	1020
	ATGGAGCCGG	ATACCAAGTA	CAAAAAGTGA	TTTCTCATCC	AAATTATGAC	TCCAAGACCA	1080
	AGAACAAATGA	CATTGCGCTG	ATGAAGCTGC	AGAAGCCTCT	GACTTTCAAC	GACCTAGTGA	1140
	AACCAGTGTG	TCTGCCCAAC	CCAGGCATGA	TGCTGCAGCC	AGAACAGCTC	TGCTGGATT	1200
65	CCGGGTGGGG	GGCCACCGAG	GAGAAAGGGA	AGACCTCAGA	AGTGTGTAAC	GCTGCCAAGG	1260
	TGCTTCTCAT	TGAGACACAG	AGATGCAACA	GCAGATATGT	CTATGACAAC	CTGATCACAC	1320
	CAGCCATGAT	CTGTGCCCGG	TTCTGCAGG	GGAACGTCGA	TTCTTGCCAG	GGTGACAGTG	1380
	GAGGGCCTCT	GGTCACTTCG	AACAACAATA	TCTGGTGGCT	GATAGGGGAT	ACAAAGCTGGG	1440
	GTCTTGCGTG	TGCCAAAGCT	TACAGACCAG	GAGTGTACGG	GAATGTGATG	GTATTACACG	1500
70	ACTGGATT	TGACAAATG	AAGGCAACG	GCTAATCCAC	ATGGTCTTCG	TCCTTGACGT	1560
	CGTTTACAA	GAAAAAATG	GGGCTGGTTT	TGCTTCCCCG	TGCATGATTT	ACTCTTAGAG	1620
	ATGATTCAGA	GGTCACTTCA	TTTTTATTA	ACAGTGAAC	TGCTTGGCTT	TGGCACTCTC	1680
	TGCCATACTG	TGCAGGCTGC	AGTGGCTCCC	CTGCCAGCC	TGCTTCCCT	AACCCCTTGT	1740
75	CCGCAAGGGG	TGATGGCCGG	CTGGTTGTGG	GCACTGGCGG	TCAATTGTGG	AAGGAAGAGG	1800
	GTGGAGGCT	GCCCCATTG	AGATCTTCCT	GCTGAGTCTC	TTCCAGGGGC	CAATTTTGGA	1860
	TGAGCATGGA	GCTGTCACTT	CTCAGCTGCT	GGATGACTTG	AGATGAAAAA	GGAGAGACAT	1920
	GGAAAGGGAG	ACAGCCAGGT	GGCACCTGCA	GCGGCTGCCC	TCTGGGGCCA	CTTGGTAGTG	1980
	TCCCCAGCCT	ACTTCAACAAG	GGGATTTTGC	TGATGGGTTT	TTAGAGCCCT	AGCAGCCCTG	2040
	GATGGTGGCC	AGAAATAAAG	GGACCAAGCC	TTCATGGGTG	GTGACGTGGT	AGTCACTTGT	2100
	AAGGGGAACA	GAAACATTTT	TGTTCTTATG	GGGTGAGAAT	ATAGACAGTG	CCCTTGGTGC	2160

GAGGGAAGCA ATTGAAAAGG AACTTGCCCT GAGCACTCCT GGTGCAGGTC TCCACCTGCA 2220
 CATTGGGTGG GCCTCCTGGG AGGGAGACTC AGCCTTCCTC CTCACTCCTCC CTGACCCTGC 2280
 TCCTAGCACC CTGGAGAGTG AATGCCCTTT GGTCCCTGGC AGGGCGCCAA GTTTGGCACC 2340
 ATGTCGGCCT CTTCAGGCCT GATAGTCATT GGAAATTGAG GTCCATGGGG GAAATCAAGG 2400
 ATGCTCAGTT TAAGGTACAC TGTTTCCATG TTATGTTTCT ACACATTGAT GGTGGTGACC 2460
 CTGAGTTCAA AGCCATCTTT

SEQ ID NO:173 PEL3 Protein sequence:

Protein Accession #: NP_005647.1

1 11 21 31 41 51
 MALNSGSPPA IGPYYENHGY QPENPYPAQP TVVPTVYEVH PAQYYPSPVP QYAPRVLTQA 60
 SNFVVCTQPK SPSGTVCCTSK TKKALCITLT LGTFLVGAAL AAGLLWKFMG SKCSNSGIEC 120
 DSSGTCINPS NWCDGVSHCP GGEDENRCVR LYGNFILLQM YSSQRKSWHP VCQDDWNENY 180
 GRAACRDMGY KNNFYSSQGI VDDSGSTFSM KLNTSAGNVD IYKKLYHSDA CSSKAVVSLR 240
 CLACGVNLNS SRQSRIVGGE SALPGAWPWQ VSLHVQNVHV CCGSIITPEW IVTAAHCVEK 300
 PLNNPWHWTA FAGILRQSFM FYGAGYQVQK VISHENYDSK TKNNIDIALMK LQKPLTFNDL 360
 VKFPVCLPNFG MMLQPEQLCW ISGWGATEEK GKTSEVLNAA KVLIIETQRC NSRYVVDNLI 420
 TPAMICAGFL QGNVDSQCGD SGGPLVTSNN NIWWLIGDTS WSGSCAKAYR PGVYGNVMVF 480
 TDWIYRQMKR NG

SEQ ID NO:174 PBJ4 DNA SEQUENCE

Nucleic Acid Accession #:

AI694767

Coding sequence:

130-1086 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 CAGAGAGGCT GTATTTTCAGT GCAGCCTGCC AGACCTCTTC TGGAGGAAGA CTGGACAAAG 60
 GGGGTACACAC ATTCCTTCCA TACGGTTGAG CCTCTACCTG CCTGGTGCTG GTCACAGTTC 120
 AGCTTCTTCA TGATGGTGGA TCCCAATGGC AATGAATCCA GTGCTACATA CTTCATCCTA 180
 ATAGGCTTCC CTGGTTTAGA AGAGGCTCAG TTCCTGGTTGG CCTTCCCATTT GTGCTCCCTC 240
 TACCTTATGT CTGTGCTAGG TAACTTGACA ATCATCTACA TTGTGCGGAC TGAGCACAGC 300
 CTGCTATGAG CCATGTATAT ATTTCTTTGC ATGCTTTTCAG GCATTGACAT CCTCATCTCC 360
 ACCTCATCCA TGCCCAAAT GCTGGCCATC TTCTGGTTCA ATTCCACTAC CATCCAGTTT 420
 GATGCTGTGC TGCTACAGAT GTTTGCCATC CACTCCTTAT CTGGCATGGA ATCCACAGTG 480
 CTGCTGGCCA TGGCTTTTGA CCGCTATGTG GCCATCTGTC ACCCACTGCG CCATGCCACA 540
 GTACTTACGT TGCTCTGTGT CACCAAAATT GGTGTGGCTG CTGTGGTGCG GGGGGCTGCA 600
 CTGATGGCAG CCCTTCTGTG CTTCATCAAG CAGCTGCCCT TCTGCCGCTC CAATATCCTT 660
 TCCCATTCCT ACTGCCCTACA CCAAGATGTC ATGAAGCTGG CCTGTGATGA TATCCGGGTC 720
 AATGTCGTCT ATGGCCTTAT CGTCATCATC TCCGCCATTG GCCTGGACTC ACTTCTCATC 780
 TCCTTCTCAT ATCTGCTTAT TCTTAAGACT GTGTGGGCT TGACACGTGA AGCCCAAGGC 840
 AAGGCATTTG GCACCTTGCGT CTCTCATGTG TGTGCTGTGT TCATATTTCTA TGTACCTTTC 900
 ATTGGATGTG CCATGGTGCA TCGCTTTAGC AAGCGGCGTG ACTCTCCACT GCCCGTCATC 960
 TTGGCCAATA TCTATCTGCT GGTTCCTCCT GTGCTCAACC CAATTGTCTA TGGAGTGAAG 1020
 ACAAGAGAGA TTGACAGAGG CATCCTTCGA CTTTTCATG TGGCCACACA CGCTTCAGAG 1080
 CCCTAGGTGT CAGTGATCAA ACTTCTTTTC CATTCAGAGT CCTCTGATTC AGATTTTAAT 1140
 GTTAACATTT TGAAGACAG TATTCAGAAA AAAAATTTCC TTAATAAAAA TACAACCTCAG 1200
 ATCCTTCAAA TATGAAACTG GTTGGGGAAT CTCCATTTTT TCAATATTAT TTTCTTCTTT 1260
 GTTTTCTTGC TACATATAAT TATTAATACC CTGACTAGGT TGTGGTTGGA GGGTTATTAC 1320
 TTTTCAATTT ACCATGCACT CCAATCTAA ACTGCTTCTA CTGATGGTTT ACAGCATTTCT 1380
 GAGATAAGAA TGGTACATCT AGAGAACAT TGCCAAAGGC CTAAGCACAG CAAAGGAAAA 1440
 TAAACACAGA ATATAATAAA ATGAGATAAT CTAGCTTAAA ACTATAACTT CCTCTTCAGA 1500
 ACTCCCAACC ACATTTGGATC TCAGAAAAAT ACTGCTTCA AAATGACTTC TACAGAGAAG 1560
 AAATAATTTT TCCTCTGGAC ACTAGCACTT AAGGGGAAGA TTGGAAGTAA AGCCTTGAAA 1620
 AGAGTACATT TACCTACGTT AATGAAAGTT GACACACTGT TCTGAGAGTT TTCACAGCAT 1680
 ATGGACCCTG TTTTCTCTAT TTAATTTTCT TATCAACCCCT TTAATTAGGC AAAGATATTA 1740
 TTAGTACCCCT CATTTGATGCC ATGGGAAAAAT TGATGTTTCAG TGGGGATCAG TGAATTAAAT 1800
 GGGGTACATC AAGTATAAAA ATTAATAAAA AAAGACTTCA TGCCCAATCT CATATGATGT 1860
 GGAAGAACTG TTAAGAGAGC CAACAGGGTA GTGGGTTAGA GATTTCCAGA GTCTTACATT 1920
 TTCTARAGGA GGTATTTAAT TTCTTCTCAC TCATCCAGTG TTGTATTTAG GAATTTCCCTG 1980
 GCAACAGAAC TCATGGCTTT AATCCCACTA GCTATTGCTT ATTGCTCTGG TCCAATTGCC 2040
 AATTACCTGT GTCTTGAAG AAGTGATTTT TAGGTTTACC ATTTATGGAAG ATTCTTATTTC 2100
 AGAAAGTCTG CATAGGGCTT ATAGCAAGTT ATTTATTTTT AAAAGTTCCA TAGGTTTTC 2160
 TGAATAGGAG TGAGGTTAGG GAGCCACCAG TTATGATGGG AAGTATGGAA TGGCAGGTGT 2220
 TGAAGATAAC ATTGGCCCTT TGAGTGTGAC TCGTAGCTGG AAAGTGAGGG AATCTTCAGG 2280
 ACCATGCTTT ATTTGGGGCT TTGTGCAGTA TGGAACAGGG ACTTTGAGAC CGGGAAAGCA 2340
 ATCTGACTTA GGCATGGGAA TCAGGCATTT TTGCTTCTGA GGGGCTATTA CCAAGGGTTA 2400
 ATAGGTTTCA TCTTCAACAG GATATGACAA CAGTCTTAAC CAAGAAACTC AAATTACATA 2460
 TACTAAAACA TGTGATCATA TATGTGGTAA GTTTCATTTT CTTTTCATAT CCTCAGGTTTC 2520
 CCTGATATGG ATTCCTATNA CATGCTTTCA TCCCTTTTTC TAATGGATAT CATATTGGGA 2580
 AATGCTTATT TAATCATTTG ATTTGCTGCT GGACTGTAAG CCCATGAGGG CACTGTTTAT 2640
 TATTTGAATG CATCTCTGTT CATCATTTGAC TGCTCTTTGC TCATCATTTGA ATCCCCCAGC 2700
 AAAGTGCCTA GAACATAATA GTGCTTATGC TTGACACCGG TTATTTTTC TCAAACTGA 2760
 TTCCTTCTGT GCTGAACACA TAGCCAGGCA ATTTTCCAGC CTCTTTTTCAG TTGGGTATTA 2820
 TTAATTTTCA GCCATTTACT CCAATGTGAG TGGAAAGTAC ATGTGCAATT TTTATACCTG 2880
 GCTCATAAAA CCGTCCCATG TGCAGCCTTT CATGTTGACA TTAATGTGTA CTTGGGAAGC 2940

TATGTGTTAC ACAGAGTTAA TTAACCNAA AGGCCTGGNA ATTTTTTGNN AANNAACTG 3000
TGGCCNNAG GCCCNCACCC CTTTTNNNA ATTTGGCAAN NTCCCACCTT GTANTTTGGT 3060
AAGGAGGCCA GTTGGAATAAG TGAAAAATAA AGTACTATTG TGTC

5

Protein Accession #: SEQ ID NO:175 PBJ4 PROTEIN SEQUENCE
not available, cloned at Eos

10

15

1	11	21	31	41	51	
MVDPNGNESS	ATYFILIGLP	GLEEAQFWLA	FPLCSLYLIA	VLGNLTIIYI	VRTEHSLHEP	60
MYIFLCMLSG	IDILISTSSM	PKMLAIFWFN	STTIQFDACL	LQMFIAHSLS	GMESTVLLAM	120
AFDRYVAICH	PLRHATVLTL	PRVTKIGVAA	VVRGAALMAP	LPVFIKQLPF	CRSNILSHSY	180
CLHQDVMKLA	CDDIRNVVYV	GLIVIIISAIG	LDSLLISFSY	LLILKTVLGL	TREAQAKAFG	240
TCVSHVCAVF	IFVYVFFIGLS	MVHRFSKRDR	SPLFVILANI	YLLVPPVLNP	IVYGVKTKEI	300
RQRILRLFHV	ATHASEP					

20

Nucleic Acid Accession #: SEQ ID NO:176 PM72 DNA SEQUENCE
Coding sequence: NM_004624.1
57-1544 (underlined sequences correspond to start and stop codons)

25

30

35

40

45

50

55

60

65

70

TCGGAGCCTG	CGGAGGGTGG	TGGTGGTGGT	GGTGGTGGCC	CTCGCCCGCC	TCACTCATGC	60
CTCCTCCTCC	TCGTCTCTCG	CTCAGGCGCC	TCGGTGGCGG	TTGGTGGCGG	GTTACGCGGC	120
TGGTGGTCCG	GGCGGCGCGG	GCTCGCTCTC	GGGGAGGCCG	GGGCGGATCT	CGCGGCGCAG	180
GGCGGCGCGG	CCGAGGTGGG	GTCCGCGCGC	GGAGGCGGCT	CGAGCTTCGT	GCTGCGCGCT	240
CGCTCTTGGG	CTCCTCGCTG	CAGGAGGAGT	GTGACTATGT	GCAGATGATC	GAGGTGCAGC	300
ACAAGCAGTG	CCTGGAGGAG	GCCCAGCTGG	AGAATGAGAC	AATAGGCTGC	AGCAAGATGT	360
GGGACAACCT	CACCTGCTGG	CCAGCCACCC	CTCGGGGCCA	GGTAGTTGTC	TTGGCCTGTC	420
CCCTCATCTT	CAAGCTCTTC	TCCTCCATTC	AAGGCCGCAA	TGTAAGCCGC	AGCTGCACCG	480
ACGAAGCGTG	GACGCACCTG	GAGCCTGGCC	CGTACCCCAT	TGCCTGTGGT	TTGGATGACA	540
AGGCAGCGAG	TTTGGATGAG	CAGCAGACCA	TGTTCTACGG	TTCTGTGAAG	ACCGGCTACA	600
CCATTGGCTA	CGGCCCTGTCC	CTCGCCACCC	TTCTGGTCCG	CACAGCTATC	CTGAGCCTGT	660
TCAGGAAGCT	CACTGCACG	CGGAATACA	TCCACATGCA	CCTCTTCATA	TCCTTCATCC	720
TGAGGGCTGC	CGCTGTCTTC	ATCAAAGACT	TGGCCCTCTT	CGACAGCGGG	GAGTCGGACC	780
AGTGCTCCGA	GGGCTCGGTG	GGCTGTAAAG	CAGCCATGGT	CTTTTCCCAA	TATTGTGTCA	840
TGGCTAACTT	CTTCTGGCTG	CTGGTGGAGG	GCCTCTACCT	GTACACCCCT	CTTGCCGTCT	900
CCTTCTTCTC	TGAGCCGAAG	TACTTCTGGG	GGTACATACT	CATCGGCTGG	GGGGTACCCA	960
GCACATTTCAC	CATGGTGTGG	ACCATCGCCA	GGATCCATTT	TGAGGATTAT	GGTCTGTCTA	1020
GGTGCTGGGA	CACCAATCAAC	TCCTCACTGT	GGTGGATCAT	AAAGGGCCCC	ATCCTCACTT	1080
CCATCTTGGT	AAACTTCATC	CTGTTTATTT	GCATCATCCG	AATCCTGCTT	CAGAAACTGC	1140
GGCCCCCAGA	TATCAGGAAG	AGTGACAGCA	GTCCATACTC	AAGGCTAGCC	AGGTCCACAC	1200
TCCTGCTGAT	CCCCCTGTTT	GGAGTACACT	ACATCATGTT	CGCCTTCTTT	CCGGACAATT	1260
TTAAGCCTGA	AGTGAAGATG	GTCTTTGAGC	TCGTCTGTGG	GTCTTTCCAG	GGTTTGTGTT	1320
TGGCTATCCT	CTACTGCTTC	CTCAATGGTG	AGGTGCAGGC	GGAGCTGAGG	CGGAAGTGGC	1380
GGCGCTGGCA	CTGCGAGGGC	GTCTGGGGCT	GGAACCCCAA	ATACCGGCAC	CCGTGCGGAG	1440
GCAGCAACGG	CGCCACGTGC	AGCACGCAGG	TTTCCATGCT	GACCCGCGTC	AGCCCAGGTG	1500
CCCGCGGCTC	CTCCAGCTTC	CAAGCCGAAG	TCTCCCTGGT	CTGACCACCA	GGATCCCAGC	1560
CCAAGCGGCC	CCTCCGCGCC	CTTCCCACTC	GCAGCAGACG	CCGGGGACAG	AGGCCTGCCC	1620
GGCGCGGCCA	GCCCCGGCCC	TGGGCTCGGA	GGCTGCCCCC	GGCCCCCTGG	TCTCTGGTCC	1680
GGACACTCCT	AGAAACCGCA	GCCCTAGAGC	CTGCCCTGGG	CGTTTCTAGC	AAGTGAGAGA	1740
GATGGGAGCT	CTCTCTCTGG	AGGATGCAGG	TGGAATCAG	TCATTAGACT	CCTCCTCCAA	1800
AGGCCCCCTA	CGCCAATCAA	GGGCAAAAAG	TCTACATAT	TTCACTCTGA	CTCTGCCCCC	1860
TGCTGGCTCT	TCTGGCCAAAT	TGGAGGAAAG	CAACCGGTGG	ATCCTCAAAC	AACACTGGTG	1920
TGACCTGAGG	GCAGAAAGGT	TCTGCCCGGG	AAGGTCACCA	GCACCAACAC	CACGGTAGTG	1980
CCTGAAATTT	CACCATTTGCT	GTCAAGTTCC	TTTGGGTTAA	GCATTACCAC	TCAGGCATTT	2040
GACTGAAGAT	GCAGCTCACT	ACCCTATTCT	CTCTTTACGC	TTAGTTATCA	GCTTTTAAAA	2100
GTGGGTATAT	CTGGAGTTT	TGTTTGGAGA	GCACACCTAT	CTTAGTGGTT	CCCCACCGAA	2160
GTGGACTGGC	CCCTGGGTCA	GTCTGGTGGG	AGGACGGTGC	AACCCAAGGA	CTGAGGGACT	2220
CTGAAGCCTC	TGGGAAATGA	GAAGGCAGCC	ACCAGCGAAT	GCTAGGTCTC	GGACTAAGCC	2280
TACCTGCTCT	CCAAGTCTCA	GTGGCTTCAT	CTGTCAAGTG	GGACTCTGTC	ACACCAGCCA	2340
TTCTTATCTC	TCTGTCTCTG	GGAAGCAACA	GGAATCAAGA	GACTGCCCTC	CTTGTCCACC	2400
CACCTATGTG	CCAATGTTG	TAACTAGGCT	CAGAGATGTG	CACCCATGGG	CTCTGACAGA	2460
AAGCAGATCC	TCACCTGTCT	ACACATACAG	GATTTGAAC	CAGATCTGTC	TGATAGGAAT	2520
GTGAAAGCAC	GGAATCTTAC	TGCTAACTTT	TGTGTATCGT	AACCCAGCAG	ATCCTCTTGG	2580
TTATTTGTTT	ACCACTTGTA	TTATTAATGC	CATTAATCCCT	GAATTTCCCT	TGCCACCCCA	2640
CCCTCCCTGG	AGTGTGGCTG	AGGAGGCTC	CATCTCATGT	ATCATCTGGA	TAGGAGCCTG	2700
CTGGTCACAG	CCTCTCTCTG	CTGCCCTTCA	CCCCAGTGGC	CACCTAGCTT	CCTACCCACA	2760
CTCTGCCAG	AGATACCCCT	CAGGACTGCA	ACAGGCTTGT	GCAACAATAA	ATGTTGGCTT	2820
GGAAAAA	AAAA					

75

SEQ ID NO:177 PM72 Protein sequence:
Protein Accession #: JC2195

80

1	11	21	31	41	51	
MPPPPLLSLR	RLGGWSAVT	RLVVAAGAR	SRGGRGSRG	AGGGGRGGVA	RRRRLRLRAA	60
RSLLGSSLQE	ECDYVQIEV	QHKQCLEEAQ	LENETIGCSK	MWNLTCWPA	TPRGQVVVLA	120

CPLIFKLFSS IQGRNVSRSC TDEGWTHLEP GYPPIACGLD DKAASLDEQQ TMFYGSVKTG 180
 YTIYGLSLA TLLVATAILS LFRKLHCTRN YIHMHLFISF ILRAAAVFIK DLALFDSGES 240
 DQCEGSESVGC KAAMVFYQYC VMANFFWLLV EGLYLYTLLA VSFFSERKYF WGYILIGWGV 300
 PSTFTMVWTI ARIHFEDYGL LRCWDTINSS LWIWKGPIL TSILVNFILF ICIIIRILLQK 360
 LRPFDIRKSD SSPYSRLARS TLLLIPLFGV HYIMFAFFPD NFKPEVKMVF ELVVGSFQGF 420
 VVAILYCFLN GEVQAELERRK WRRWHLQGV L GWNPKYRHPS GGSNGATCST QVSMLTRVSP 480
 GARRSSSFQA EVSLV

SEQ ID NO:178 BFF8 DNA SEQUENCE

Nucleic Acid Accession #: AL133619
 Coding sequence: 1-2070 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 15 ATGAGCGGTG CGGGGGTGGC GGCTGGGACG CGGCCCCCCA GCTCGCCGAC CCCGGGCTCT 60
 CGGCGCCGGC GCCAGCGCCC CTCTGTGGGC GTCCAGTCTT TGAGGCCGCA GAGCCCGCAG 120
 CTCAGGCAGA GCACCCGCGA GAAACGGAAC CTGGACCTGG AGAAAAGCCT GCAGTTCCTG 180
 AGCAGCAGC ACTCGGAGAT GCTGGCCAAG CTCCATGAGG AGATCGAGCA TCTGAAGCGG 240
 20 GAAAAACAAG GTGAGCCGGC GCGGGGCCCT AGGCCGGCCC TGCCCTCCCA GGCACACTCA 300
 ACCTGCGCC TCCCGCAGCA CAGAAACACA GCCATCAACT CCAGCACACG CCTGGGCTCA 360
 GGGGGAACAC AGGACGGGGA GCGCCCTCCG ACTGTCTTTC CCCACCTGGC TGCCTGGGCC 420
 CCTGTATGCC AACCCAGTGG GTACAGGTTC TGGGGGACCT GGACAGATGC CGCTACCTCT 480
 AGCCGTGGCT GGACGATGTT ATGCAGCCAA GCACAGCAGC TGCTGCTCTC GGGAAGCCCA 540
 GGGCTTGAGG TCATTGCAGG GCGGCAGGTG GCCACAGGGT GCTCCCCAGA CCTCCCTCCT 600
 25 CCAAGTAGAG CTGAAATGGG AAGGAACCCC TGGGACAGCC CCTGCGCTGC TAGATCTTTG 660
 CCTCAGATTG CTGCTGTGGC CAGGCCCAGG ATTTCCAGCC CTATGGCTCT GAGTCTCTAC 720
 ATGCTGGGGG CCCAGGGGAT ATGGACACAC TCCATCCAGG GATCCCTTCC TGCCATCTGG 780
 GCAGCAACCA TGGGGACAAA GGGAGGAAGC AGAGTCCTGT TTCCTTGCCA CTTGTCCAAG 840
 30 GCACTTCCCC ATCTGCAGC GCGCCCCAC CCAGCCCAGG ATCCTGGGCT GTGGTCTCAA 900
 GCTCACTTCC CATTATCTTT GGGGCTGGGG CTGACATCAG GAGGACATCT GACTGGTGA 960
 TGGAGCCAGC CTGGGAACAT CGCAGCTGGG GCAGTGCCCTA GGGCTCTCCC TTCCAGGGA 1020
 GACATGGAGA AGGGGTTGA GGGAGGGCCC TTCCCTAGCC GCTGTGGCAA CTCCAGTGAG 1080
 CTGTTCTGGG CAAAGTGTGG CCCAAGTCGG CAGCCCCAGC CCTGCAGTGC TGGGGACGCT 1140
 35 GACAGGACAC GGAAGAGGCG CATGCTTTCC CTGCGGACCT GCTGTTCCTAT GTGTCCCAAG 1200
 CCTCTCTGCT TTCCAGATGG CCCTCAGGA AACCACTTTT CCAGGGCCTC TGCTCCCTTG 1260
 GCGCTCTGCT GGTCTTCGAT CAACGGAGTG TGGGTAGAGC CGGGAGGACC CAGCCCTGCC 1320
 AGGCTGAAGG AGGGCTCCTC ACGGACACAC AGGCCAGGAG GCAAGCGTGG GCGTCTTGGC 1380
 GCGGTAGCG CCGACACTGT GCGCTCTCCT GCAGACAGCC TCTCCATGTC AAGCTTCCAG 1440
 40 TCTGTCAAGT CCATCTCTAA TTCAGCCAAC TCTCAAGGCA AGGCCAGGCC CCAGCCCGGC 1500
 TCCTTCAACA AGCAAGATTC AAAAGCTGAC GTCTCCCAAG AGGCCGACCT GGAAGAGGAG 1560
 CCCTACTTTC ACAACAGCAA GCTGGACAAA GTTCTTGGGG TACAAGGGCA GGCCAGAAAG 1620
 GAGAAAGCAG AGGCCTCTAA TGCAGGAGCT GCCTGTATGG GGAACAGCCA GCACCAGGGC 1680
 AGGCAGATGG GGGCGGGGGC ACACCCCCCA ATGATCCTGC CCCTTCCCTT GCGAAAGCCC 1740
 45 ACCACACTTA GGCAGTGGCA AGTGCTCATC CGCGAGCTGT GGAATACCAA CCTCTGCGAG 1800
 ACCCAAGAGC TGGGGCACCT CAAGTCCCTC CTGGAAGGGA GCCAGAGGCC CCAGGCAGCC 1860
 CCGGAGGAAG CTAGCTTTTC CAGGGACCAA GAAGCCACGC ATTTCCCCAA GGTCTCCACC 1920
 AAGAGCCTCT CCAAGAAATG CCTGAGCCCA CCTGTGGCGG AGCGTGCCAT CCTGCCCGCA 1980
 CTGAAGCAGA CCCCAGAA CAACCTTGCC GAGAGGCAGA AGAGGCTGCA GGCAATGCAG 2040
 50 AAACGGCGCC TGCATCGCTC AGTGCTTTGA

SEQ ID NO:179 BFF8 Protein sequence:

Protein Accession #: T43457

55 1 11 21 31 41 51
 MSGAGVAAGT RPPSSPTPGS RRRRQRPVSV VQSLRPQSPQ LRQSDPQKRN LDLEKSLQFL 60
 QQQHSEMLAK LHEEIEHLKR ENKGEPAARG RPALPPQAH S TLPLPQHRNT AINSSTRLGS 120
 60 GGTQDGEPLQ TVLAHLAALA PVCQPSGYRF WGTWTDAA TS SRGWTMLCSQ AQHVLLSGSP 180
 GPEVIAGRQV ATGCSDDLFP PSRAEMGRNP WDSPCPARSL PQIAA VARPR ISSPMALSPH 240
 MLGAQGIWTH SIQGS LPAIW AATMGTKGGS RVLFPCHLSK ALPHPD SGPH PAQDPGLWSQ 300
 AHFPLSLGLG L TSGGHLTGG WSQPGNIAAG AVPRALPSQG DMEKGVGGP FFSRCGNSS 360
 LFWAKCGPSR QPQPCAGDA DRTEEAMLS LGTCCSMCPK PSCFPDGP SG NHLSRASAPL 420
 65 GARWVINGV WVEPGGSPSA RLKESGSRTH RPPGKRGRLA GGSADTVRSP ADSLSMSSFQ 480
 SVKISNSAN SQKARFPQG SFNKQDSKAD VSQKADLEEE PLLHNSKLDK VFGVQGQARK 540
 EKAEASNAGA ACMGNSHQHG RQMGAGAHPP MILPLPLRKP TTLRQCEVLI RELWNTNLLQ 600
 TQELRHLKSL LEGSQRFQAA PEEASFPRDQ EATHFFKVST KSLSKKCLSP PVAERAILPA 660
 LKQTPKNFNA ERQKRLQAMQ KRRLHRSVL

SEQ ID NO:180 BCR4 DNA SEQUENCE

Nucleic Acid Accession #: NM_012319.2
 Coding sequence: 138-2405 (underlined sequences correspond to start and stop codons)

75 1 11 21 31 41 51
 CTCGTGCCGA ATTCGGCACG AGACCCGCTG TTCGCGCCTG GTAGAGATTT CTCGAAGACA 60
 CCAGTGGGCC CGTGTGGAAC CAAACCTGCG CGCGTGGCCG GGCCGTGGGA CAACGAGGCC 120

5
10
15
20
25
30
35
40
45
50
55

CGGGAGACGA AGGCGCAATG GCGAGGAAGT TATCTGTAAT CTTGATCCTG ACCTTTGCCC 180
TCTCTGTAC AAATCCCTT CATGAACATA AAGCAGCTGC TTTCCCCAG ACCACTGAGA 240
AAATTAGTCC GAATTGGGAA TCTGGCATTG ATGTTGACTT GGCAATTTCC ACACGGCAAT 300
ATCATCTACA ACAGCTTTTC TACCGCTATG GAGAAAATAA TTCTTTGTCA GTTGAAGGGT 360
TCAGAAAATT ACTTCAAAAT ATAGGCATAG ATAAGATTAA AAGAATCCAT ATACACCATG 420
ACCACGACCA TCACTCAGAC CACGAGCATC ACTCAGACCA TGAGCGTCAC TCAGACCATG 480
AGCATCACTC AGACCACGAG CATCACTCTG ACCATGATCA TCACCTCTAC CATAATCATG 540
CTGCTTCTGG TAAAAATAAG CGAAAAGCTC TTTGCCCAGA CCATGACTCA GATAGTTTCA 600
GTAAAGATCC TAGAAACAGC CAGGGGAAAG GAGCTCACC ACCAGAACAT GCCAGTGGTA 660
GAAGGAATGT CAAGGACAGT GTTAGTGCTA GTGAAGTGAC CTCAACTGTG TACAACATG 720
TCTCTGAAGG AACTCACTTT CTAGAGACAA TAGAGACTCC AAGACCTGGA AAACCTTTCC 780
CCAAAGATGT AAGCAGCTCC ACTCCACCCA GTGTCACATC AAAGAGCCGG GTGAGCCGGC 840
TGCTGGTATG GAAAAAATAA GAATCTGTGA GTGAGCCCG AAAAGGCTTT ATGTATTCCA 900
GAAACACAAA TGAAAAATCT CAGGAGTGT TCAATGCATC AAAGCTACTG ACATCTCATG 960
GCATGGGCAT CCAGGTTCCG CTGAATGCAA CAGAGTTCAA CTATCTCTGT CCAGCCATCA 1020
TCAACCAAAAT TGATGCTAGA TCTTGTCTGA TTCATACAAG TGAAAAGAAG GCTGAAATCC 1080
CTCCAAAGAC CPATTCATTA CAAATAGCCT GGGTTGGTGG TTTTATAGCC ATTTCCATCA 1140
TCAGTTTCTT GTCTCTGTCT GGGGTATCT TAGTGCTCT CATGAATCGG GTGTTTTTCA 1200
AAATTCTCCT GAGTTTCTGG GTGGCACTGG CCGTTGGGAC TTTGAGTGGT GATGCTTTTT 1260
TACACCTTCT TCCACATTCT CATGCAAGTC ACCACCATAG TCATAGCCAT GAAGAACCAG 1320
CAATGGAAAT GAAAAGAGGA CCACCTTTTCA GTCATCTGTC TTCTCAAAAC ATAGAAGAAA 1380
GTGCTTATTT TGATTTCCAG TGGAAAGGTC TAACAGCTCT AGGAGGCCCTG TATTTTATGT 1440
TTCTTGTGTA ACATGTCTCT ACATTGATCA AACAATTTAA AGATAAGAAG AAAAAAATC 1500
AGAAGAAACC TGAAAATGAT GATGATGTGG AGATTAAGAA GCAGTTGTCC AAGTATGAAT 1560
TCAAACTTTC AACAAATGAG GAGAAAGTAG ATACAGATGA TCGAAGTCAA GGCTATTTAC 1620
GAGCAGACTC ACAAGAGCCC TCCCACTTTG ATCTCTAGCA GCCTGCAGTC TTGGAAGAAG 1680
AAGAGTCAAT GATAGTCTAT GCTCATCCAC AGGAAGTCTA CAATGAATAT GTACCCAGAG 1740
GGTGCAAGAA TAAATGCCAT TCACATTTCC ACGATACACT CGGCCAGTCA GACGATCTCA 1800
TTCACCACCA TCATGACTAC CATCATATTC TCCATCATCA CCACCACCAA AACACCATTC 1860
CTCAGAGTCA CAGCCAGCGC TACTCTCGGG AGGAGCTGAA AGATGCCGGC GTCGCCACTT 1920
TGGCTTGGAT GGTGATAATG GGTGATGGCC TGCACAAATT CAGCGATGGC CTAGCAATTG 1980
GTGCTGCTTT TACTGAAGGC TTATCAAGTG GTTTAAGTAC TTCTGTGCT GTGTTCTGTC 2040
ATGAGTTGCC TCATGAATTA GGTGACTTTG CTGTTCTACT AAAGGCTGGC ATGACCGTTA 2100
AGCAGGCTGT CCTTTATAAT GCATTGTCTG CCATGCTGGC GTATCTTGGG ATGGCAACAG 2160
GAATTTTCAT TGGTCATTAAT GCTGAAAATG TTCTATGTG GATATTTGCA CTTACTGCTG 2220
GCTTATTCAT GTATGTGCT CTGGTTGATA TGGTACCTGA AATGCTGCAC AATGATGCTA 2280
GTGACCATGG ATGTAGCCGC TGGGGGTAT TCTTTTACA GAATGCTGGG ATGCTTTTGG 2340
GTTTGGAAAT TATGTTACTT ATTTCCATAT TTGAACATAA AATCGTGGTT CGTATAAATT 2400
TCTAGTTAAG GTTTAAATGC TAGAGTAGCT TAAAAGTTG TCATAGTTTC AGTAGGTCT 2460
AGGAGATAG GTTTGTATGC TGTACTATGC AGCGTTTAAA GTTAGTGGGT TTTGTGATTT 2520
TTGTATTGAA TATGCTGTCT TGTACAAAG TCAGTTAAAG GTACGTTTTC ATATTTAAGT 2580
TATTTCTATCT TGGAGATAAA ATCTGTATGT GCAATTCACC GGTATTACCA GTTTATTATG 2640
TAAACAAGAG ATTTGGCATG ACATGTTCTG TATGTTTCAG GGAATAATGT CTTTAATGCT 2700
TTTTCAAGAA CTAACACAGT TATTCCTATA CTGGATTTTA GGTCTCTGAA GAACTGTCTG 2760
TGTTTAGGAA TAAGATGTG CATGAAGCCT AAAATACCAA GAAAGCTTAT ACTGAATTTA 2820
AGCAAGAGAA TAAAGAGAA AAGAGAGAA TCTGAGAAAT GGGGAGGCAT AGATTCTTAT 2880
AAAAATCACA AAATTTGTTG TAAATTAGAG GGGAGAAATT TAGAATTAAG TATAAAAGG 2940
CAGAATTAGT ATAGAGTACA TTCATTAAAC ATTTTGTCA GGATATTTC CCGTAAAAAC 3000
GTAGTGAGCA CTCTCATATA CTAATTAGTG TACATTTAAC TTTGTATAAT ACAGAAATCT 3060
AAATATATTT AATGAATTCA AGCAATATAC ACTTGACCAA GAAATTTGAA TTTCAAAATG 3120
TTCGTGCGGG TTATATACCA GATGAGTACA GTGAGTAGTT TATGTATCAC CAGACTGGGT 3180
TATTGCCAAG TTATATATCA CAAAAGCTG TATGACTGGA TGTCTGTGTT ACCTGGTTTA 3240
CAAAATATC AGAGTAGTAA AACTTTGATA TATATGAGGA TATTAAACT ACATAAGTA 3300
TCATTTGATT CGATTAGAAA AGTACTTTGA TATCTCTCAG TGCTTCAGTG CTATCATGT 3360
GAGCAATTGT CTTTATATAC GGTACTGTAG CCATCTAGG CCTGTCTGTG GCATCTCTCA 3420
GATGTTTCTT TTTTACACAA TAAATCCCTT ATATCAGCTT G

60

Protein Accession #: NP_036451

SEQ ID NO:181 BCR4 PROTEIN SEQUENCE

1 11 21 31 41 51
| | | | |
65 MARKLSVILI LTFALSVTNP LHELKAAAFP QTTEKISPWW ESGINVDLAI STRQYHLQQL 60
FYRYGENNSL SVEGFRKLLQ NIGIDKIKRI HIHHDHDHHS DHEHSDHER HSDHEHSDH 120
EHSHDHDHHS HNNHAASGN KRKALCPDHD SDSSGKDPNN SQGKAHRPE HASGRRNVKD 180
SVSASEVTST VNTVSEGHF FLETIETPRE GLKLPKDVSS STPPSVTSKS RVSLRAGRKT 240
NESVSEPRKG FMYSRNTNEN PQECFNASKL LSHGMGIQV PLNATEFNYL CPAINQIDA 300
70 RSCLHTSEK KAEIPPKTYS LQIAWVGFI AISIISFLSL LGVILVPLMN RVFFKFLLSF 360
LVALAVGTLS GDAFLHLLPH SHASHHSHS HEEPAMEMKR GPLFSLHLSQ NIEESAYFDS 420
TWKGLTALGG LYFMFLVEHV LTLIKQFKDK KKNQKKPEN DDDVEIKKQL SKYESQLSTN 480
EEKVDVDDRT EGYLRADSQE PSHFDSQOPA VLEEEVMIQ HAHPOEVYNE YVPRGCKNKC 540
75 HSHFHDTLGQ SDDLHSHHH YHHLHHHH QNHHPHSHSQ RYSREELKDA GVATLAWMVI 600
MGDGLHNFSD GLAIGAAFFE GLSSGLSTSV AVFCHELPE LGDFAVLLKA GMTVKQAVLY 660
NALSAMLAYL GMATGIFIGH YAENVSMWIF ALTAGLFMYV ALVDMVPEML HNDASDHGCS 720
RWGYFFLQNA GMLLGFGIML LISIFEHKIV FRINF

SEQ ID NO:182 BCY2 DNA sequence

Nucleic Acid Accession #:

NM_001203

Coding sequence:

274-1782 (underlined sequences correspond to start and stop codons)

5 1 11 21 31 41 51
 | | | | |
 CGCGGGGCGC GGAGTCGGCG GGGCCTCGCG GGACGCGGGC AGTGCGGAGA CCGCGGCGCT 60
 GAGGACGCGG GAGCCGGGAG CGCACGCGCG GGGTGGAGTT CAGCCTACTC TTTCTTAGAT 120
 10 GTGAAAGGAA AGGAAGATCA TTTCATGCCT TGTGTATAAA GGTCAGACT TCTGCTGATT 180
 CATAACCAATT TGGCTCTGAG CTATGACAAG AGAGGAAACA AAAAGTTAAA CTTACAAGCC 240
 TGCCATAAGT GAGAAGCAAA CTTCCTTGAT AACATGCTTT TGCGAAGTGC AGGAAAAATTA 300
 AATGTGGGCA CCAAGAAAGA GGATGGTGAG AGTACAGCCC CCACCCCCG TCCAAAGGTC 360
 TTGCGTTGTA AATGCCACCA CCATTGTCCA GAAGACTCAG TCAACAATAT TTGCAGCACA 420
 15 GACGGATATT GTTTCACGAT GATAGAAGAG GATGACTCTG GGTGCGCTGT GGTCACTTCT 480
 GGTTCCTAG CAGTAGAAGG CTCAGATTTT CAGTGTGCGG ACATCCCCAT TCCTCATCAA 540
 AGAAGATCAA TTGAATGCTG CACAGAAAGG AACGAATGTA ATAAAGACCT ACACCCTACA 600
 CTGCTCCAT TGAAAAACAG AGATTTTGTG GATGGACCTA TACACCACAG GGCTTTACTT 660
 ATATCTGTGA CTGTCTGTAG TTGTCTCTTG GTCCTTATCA TATTATTTTG TTACTTCCGG 720
 20 TATAAAGAC AAGAAACCCG ACCTCGATAC AGCATTGGGT TAGAACAGGA TGAAACTTAC 780
 ATTCTCTCTG GAGAATCCCT GAGAGACTTA ATTGAGCAGT CTCAGAGCTC AGGAAGTGGA 840
 TCAGGCCTCC CTCTGCTGGT CCAAAGGACT ATAGCTAAGC AGATTCAGAT GGTGAAACAG 900
 ATTGGAAAAG GTCGCTATGG GGAAGTTTGG ATGGGAAAGT GCGGTGCGCA AAAGGTAGCT 960
 GTGAAAGTGT TCTTACCAC AGAGGAAGCC AGCTGGTTCA GAGAGACAGA AATATATCAG 1020
 25 ACAGTGTGTA TGAGGCATGA AACATTTTG GGTTCATATG CTGCAGATAT CAAAGGGACA 1080
 GGGTCTCTGA CCCATTGTA CCTAATCACA GACTATCATG AAAATGGTTC CTTTATGAT 1140
 TATCTGAAGT CCACCACCT AGACGCTAAA TCAATGTGTA AGTTAGCCTA CTCTCTGTG 1200
 AGTGGCTTAT TCATTATACA CACAGAAATC TTAGTACTC AAGGCAAAACC AGCAATTGCC 1260
 CATCGAGATC TGAAGAAGTAA AACATTCTG GTGAAGAAAA ATGGAAGTTC CTGTATTGCT 1320
 GACCTGGGCC TGGCTGTAA ATTTATTAGT GATACAAATG AAGTTGACAT ACCACCTAAC 1380
 30 ACTCGAGTTG GCACCAAACG CTATATGCCT CCAGAAGTGT TGGACGAGAG CTGTAACAGA 1440
 AATCACTTCC AGTCTTACAT CATGGCTGAC ATGTATAGTT TTGGCCTCAT CTTTGGGAG 1500
 GTTGCTAGGA GATGTGTATC AGGAGGTATA GTGAAGAAT ACCAGCTTCC TTATCATGAC 1560
 CTAGTGCCCA GTGACCCCTC TTATGAGGAC ATGAGGGAGA TTGTGTGCAT CAAGAAGTTA 1620
 35 CGCCCTCAT TCCCAAACCG GTGGAGCAGT GATGAGTGT TAAGGCAGAT GGGAAAACCT 1680
 ATGACAGAAT CTGGGGCTCA CAATCCTGCA TCAAGGCTGA CAGCCCTGCG GGTTAAGAAA 1740
 ACACITGCCA AAATGTCAGA GTCCAGGAC ATAAACTCT GATAGGAGAG GAAAAGTAAG 1800
 CATCTCTGCA GAAAGCCAAC AGGTACTCTT CTGTTTGTGG GCAGAGCAAA AGACATCAA 1860
 TAAGCATCCA CAGTACAAGC CTTGAACATC GTCTGTCTC CCACTGGGTT CAGACCTCAC 1920
 40 CTTTCAGGGA GCGACCTGGG CAAAGACAGA GAAGCTCCA GAAGGAGAGA TTGATCCGTG 1980
 TCTGTTTGA GCGGAGAAAA CCGTTGGGTA ACTTGTTCAA GATATGATGC AT

SEQ ID NO:183 BCY2 Protein sequence

Protein Accession #:

NP_001194

45 1 11 21 31 41 51
 | | | | |
 MLLRSAGKLN VGTKKEDGES TAPTPRPKVL RCKCHHHCPE DSVNNICSTD GYCFMIEED 60
 DSGLPVVTSG CLGLESDFFQ CRDTPHPQR RSIECCTERN ECNKDLHPTL PPLKNRDFVD 120
 50 GPIHHRALLI SVTVCSLLL LILFCYFRY KRQETRPYS IGLEQDETYI PPGESLRDLI 180
 EQSQSSGSGS GLPLLQRTI AKQIQMVKQI GKGRYGEVWM GKWRGEKVAV KVFFTTEAS 240
 WFRETEIYQT VLMRHENILG FIAADIKGTG SWTQLYLITD YHENGSLYDY LKSTILDAKS 300
 MLKLAYSSVS GLCHLHTEIF STQKPAIAH RDLKSKNILV KKNGTCCIA LGLAVKFISD 360
 55 TNEVDIPFNT RVGTRKRYMPF EVLDESLNRN HFQSYIMADM YSFLILWEV ARRCVSGGIV 420
 EEYQLPYHDL VPSDFSYEDM REIVCIKKLR PSFPNRWSSD ECLRQMGKLM TECWAHPNAS 480
 RLTAIRVKKT LAKMSESDI KL

SEQ ID NO:184 CBF9 DNA sequence

Nucleic Acid Accession #:

AC005383

Coding Sequence:

328-2751 (underlined sequences correspond to start and stop codons)

65 1 11 21 31 41 51
 | | | | |
 GACAGTGTTC GCGGCTGCAC CGCTCGGAGG CTGGGTGACC CGCGTAGAAG TGAAGTACTT 60
 TTTTATTTTC AGACCTGGGC CGATGCCGCT TTAATAAACG CGAGGGGCTC TATGCACCTC 120
 CCTGGCGGTA GTTCCCTCCG CTTACAGCCG GTCCGGTCTG GCCGCCCTCT CCCAGGAGAG 180
 70 ACAACACAGT GTCCACGTT GCAGCCGCGC CCGGGGCGCC CTTCTGTGA TCCCGTAGCG 240
 CCCCTGGGCC CGAGCCGCGC CCGGGTCTGT GAGTAGAGCC GCCCGGGCAC CGAGCGCTGG 300
 TCGCCGCTCT CTTCCGTTA TATCAACATG CCCCCTTTC TGTGTCTGGA GGCCGTCTGT 360
 GTTTTCTCTG TTTCCAGAGT GCCCCATCT CTCCCTCTCC AGGAAGTCCA TGTAAAGCAA 420
 75 GAAACCATCG GGAAGATTTC AGCTGCCAGC AAAATGATGT GGTGCTCGGC TGCAGTGGAC 480
 ATCATGTTTC TGTAGATGG GTCTAACAGC GTCGGGAAAG GGAGCTTTGA AAGGTCCAAG 540
 CACTTTGCCA TCACAGTCTG TGACGGTCTG GACATCAGCC CCGAGAGGGT CAGAGTGGGA 600
 GCATTCCAGT TCAGTTCAC TCCTCATCTG GAATTCCTCT TGATTTCATT TTCAACCCAA 660

5
0
5
10
15
20
25
30
35
40
45

CAGGAAGTGA AGGCAAGAAT CAAGAGGATG GTTTCAAAG GAGGGCGCAC GGAGACGGAA 720
 CTGCTCTGA AATACCTTCT GCACAGAGGG TTGCTGGAG GCAGAAATGC TTCGTGCCC 780
 CAGATCCTCA TCATCGTCAC TGATGGGAAG TCCAGGGGG ATGTGGCACT GCCATCCAAG 840
 CAGCTGAAGG AAAGGGGTGT CACTGTGTTT GCTGTGGGG TCAGGTTTCC CAGGTGGGAG 900
 GAGCTGCATG CACTGGCCAG CGAGCCTAGA GGGCAGCAG TGCTGTGGC TGAGCAGGTG 960
 GAGGATGCCA CCAACGGCCT CPTCAGCACC CTCAGCAGCT CGGCCATCTG CTCCAGCGCC 1020
 ACGCCAGACT GCAGGTCGA GGCTCACCCC TGTGAGCACA GGACGCTGGA GATGGTCCGG 1080
 GAGTTCGCTG GCAATGCCCC ATGCTGGAGA GGATCGCGGC GGACCTTGC GGTGCTGGCT 1140
 GCACACTGTC CCTTCTACAG CTGGAAGAGA GTGTTCTTAA CCCACCTGC CACCTGCTAC 1200
 AGGACCACCT GCCCAGGCCC CTGTGACTCG CAGCCCTGCC AGAATGGAGG CACATGTGTT 1260
 CCAGAAGGAC TGGACGGCTA CCAGTGCCTC TGCCCGCTGG CCTTTGGAGG GGAGGCTAAC 1320
 TGTGCCCTGA AGCTGAGCCT GGAATGCAGG GTCGACCTCC TCTTCTGTCT GGACAGCTCT 1380
 GCGGGCACC A CTCTGGACGG CTTCCTGCGG GCCAAAGTCT TCGTGAAGCG GTTTGTGCGG 1440
 GCGGTGCTGA GCGAGGACTC TCGGGCCCGA GTGGGTGTGG CCACATACAG CAGGGAGCTG 1500
 CTGGTGGCGG TGCTGTGGG GGAGTGCCTG ATGTGCTG ACCTGGTCTG GAGCCTCGAT 1560
 GGCAATTCCTC TCCGTGGTGG CCCCACCTCG ACGGGCAGTG CCTTGGCGCA GGCGGCAGAG 1620
 CTGGGCTTCG GGAGCCGAC CAGGACCGGC CACGTAGAGT GGTGGTTTGG 1680
 CTCACGTAGT CACACTCCGA GGATGAGGTT GCGGGCCAG CGCGTCACGC AAGGGCGCGA 1740
 GAGCTGCTCC TGCTGGGTGT AGGCAGTGAG GCGGTGCGGG CAGAGCTGGA GGAGATCACA 1800
 GGCAGCCCAA AGCATGTGAT GGTCTACTCG GATCCTCAGG ATCTGTTCAA CCAATCCCT 1860
 GAGCTGCAGG GGAAGCTGTG CAGCCGGCAG CGGCCAGGGT CCGGACACA AGCCCTGGAC 1920
 CTCGTCTTCA TGTGAGACAC CTCCTGCTCA GTAGGGCCCG AGAATTTTGC TCAGATGCAG 1980
 AGCTTTGTGA GAAGCTGTGC CCTCCAGTTT GAGGTGAACC CTGACGTGAC ACAGGTCCGC 2040
 CTGGTGGTGT ATGGCAGCCA GGTGCAGACT GCCTTCGGGC TGGACACCAA ACCCAGCCGG 2100
 GCTGCGATGC TCGGGGCCAT TAGCCAGGCC CCTACCTAG GTGGGGTGGG CTCAGCCGGC 2160
 ACCGCCCTGC TGCACATCTA TGACAAAGTG ATGACCGTCC AGAGGGGTGC CCGGCTGGT 2220
 GTCCCCAAG CTGTGGTGGT GCTCACAGGC GGGAGAGGCG CAGAGGATGC AGCCGTTCT 2280
 GCCAGAAGC TGAGGAACAA TGGCATCTCT GTCTTGGTCG TGGGCGTGGG GCCTGTCTTA 2340
 AGTGAGGGTC TCGGAGGCTC TGCAGTCCC CGGGATTTCC TGATCCACGT GGCAGCTTAC 2400
 GCCGACCTGC GGTACACCA GACGTGCTC ATTGAGTGGC TGTGTGGAGA AGCCAAGCAG 2460
 CCAGTCAACC TCTGCAAAACC CAGCCCGTGC ATGAATGAGG GCAGCTGCGT CCTGCAGAA 2520
 GGGAGCTACC GCTGCAAGTG TCGGGATGGC TGGGAGGGCC CCCACTGCGA GAACCGTGAG 2580
 TGGAGCTCTT GCTCTGTATG TGTGAGCCAG GGATGGATTC TTGAGACGCC CCGAGGCAC 2640
 ATGGCTCCCG TGCAGGAGGG CAGCAGCCGT ACCCTTCCCA GCAACTACAG AGAAGGCCGT 2700
 GGCACGTAAA TGGTGCTTAC CTCTTGAAT GTCTGTGCCC CAGGTCTTGA GAATGTCTGC 2760
 TTCCCGCCGT GGCCAGGACC ACTATTCTCA CTGAGGGAGG AGGATGTCCC AACTGCAGCC 2820
 ATGCTGCTTA GAGACAAGAA AGCAGCTGAT GTCACCCACA AACGATGTTG TTGAAAAGTT 2880
 TTGATGTGTA AGTAAATACC CACTTCTGTG ACCTGCTGTG CCTTGTGAG GCTATGCTAT 2940
 CTGCCACCTT TCCCTTGAGG ATAAACAAGG GTCTCTGAAG ACTTAAATTT AGCGCCCTGA 3000
 CGTTCCTTG CACACAATCA ATGCTCGCCA GAATGTTGTT GACACAGTAA TGCCAGCAG 3060
 AGGCTTTATG TAGAGCATCC TTTGGACCGC GAAGGCCACG GCCTTTCAAG ATGGAAAGCA 3120
 GCAGCTTTTC CACTTCCCCA GAGACATTCT GGATGCATTT GCATTGAGTC TGAAAAGGGG 3180
 CTTGAGGGAC GTTGTGACT TCTTGGCGAC TGCCTTTGTG GTGTGGAAGA GACTTGAAGA 3240
 GGTCTCAGAC TGAATGTGAC CAATTAACCA GCTTGGTTGA TGATGGGGGA GGGGCTGAGT 3300
 TGTGCATGGG CCCAGTCTG GAGGGCCACG TAAATCGTT CTGAGTCGTG AGCAGTGTCC 3360
 ACCTTGAAGG TCTTC

SEQ ID NO:185 CBF9 Protein sequence
 Protein Accession #: none found

50

1 11 21 31 41 51
 | | | | |
 55 MPPFLLEAV CVFLFSRVPP SLPLQEVHVS KETIGKISAA SKMMWCSAAV DIMFLLDGSN 60
 SVKGGSFERS KHFAITVCDG LDISPERVRV GAFQFSSTPH LEFPLDSFST QQEVKARIKR 120
 MVFKGGRRET ELALKYLLHR GLPGGRNASV PQILLIIVTDG KSQGDVALPS KQLKERVTV 180
 FAVGVRFPRW EELHALASEP RGQHVLLAEQ VEDATNGLFS TLSSSAICSS ATPDCRVEAH 240
 PCEHRTLEMV REFAGNAFCW RGSRRTLAVL AAHC PFYSWK RVFLTHPATC YRTTCPGPCD 300
 SQPCQNGGTC VPEGLDGYQC LCPLAFGGEA NCALKLSLEC RVDLLFLDS SAGTTLDGFL 360
 50 RAKVFVKRFV RAVLSEDSRA RVGVATYSRE LLVAVPVEY QDVPDLVWSL DGIPFRGGPT 420
 LTGSALRQAA ERGFGSATRT GQDRPRRVVV LLETSHSEDE VAGPARHARA RELLLLGVS 480
 EAVRAELEEI TGSPKHVMVY SDFQDLFNQI PELQGLCSR QRPGRCTQAL DLVFMLD TSA 540
 SVGPENFAQM QSFVRSALQ FEVNPVDTQV GLVVYGSQVQ TAFGLDTPKT RAAMLRAISQ 600
 65 APYLGGVGS A GTALLHYDK VMTVQRGARP GVPKAVVLT GGRGAEDAAV PAQKLNNNGI 660
 SVLVGVGPV LSEGLRRLAG PRDSLIHVAA YADLRHQDV LIEWLCGEAK QPVNLCKPSP 720
 CMNEGSCVLQ NGSYRCKRD GWEGPHCENR EWSSCSVCVS QGWILETPLR HMAFVQEGSS 780
 RTPPSNYREG LGTEMVPTFW NVCAPEG

70

SEQ ID NO:186 PAV1 DNA sequence

Nucleic Acid Accession #: AF272890
 Coding Sequence: 87-1520 (underlined sequences correspond to start and stop codons)

75

1 11 21 31 41 51
 | | | | |
 TGCTACCCGC GCCCGGGCTT CTGGGGTGTG CCCCACCCAC GGCCAGCCCG TGCCACACCC 60
 CCCGCCCCCG GCCTCCGCG CTCGGCATGG CGCGGGGGT GCTCGTCTTG GCGCCCTCCG 120
 AGCCCGGTAA CCTGTCGTC GCCGCACCGC TCCCCGACCG CGCGGCCACC CGCGCGCGGC 180

5
10
15
20
25

```

TGCTGGTGCC CGCGTCGCGG CCCGCCTCGT TGCTGCCTCC CGCCAGCGAA AGCCCCGAGC 240
CGCTGTCTCA GCAGTGGACA GCGGGCATGG GTCTGCTGAT GCGGCTCATT GTGCTGCTCA 300
TCGTGGCGGG CAATGTGCTG GTGATCGTGG CCATCGCCAA GACGCCGCGG CTGCAGACGC 360
TCACCAACCT CTTTCATCATG TCCCTGGCCA GCGCCGACCT GGTTCATGGGG CTGCTGGTGG 420
TGCCGTTTCGG GGCACCATC GTGGTGTGGG GCGGCTGGGA GTACGGCTCC TTCTTCTGCG 480
AGCTGTGGAC CTCAGTGGAC GTGCTGTGCG TGACGGCCAG CATCGAGACC CTGTGTGTCA 540
TTGCCCTGGA CCGCTACCTC GCCATCACC CTGCCCTTCG CTACCAGAGC CTGCTGACGC 600
GCGCGCGGGC GCGGGGGCTC GTGTGCACCG TGTGGGCCAT CTCGGGCCCTG GTGTCCTTCC 660
TGCCCATCCT CATGCACTGG TGGCGGGCGG AGAGCGACGA GCGCGCGCCG TGCTACAACG 720
ACCCCAAGTG CTGCGACTTC GTACCAACCC GGGCCTACGC CATCGCCTCG TCCGTAGTCT 780
CCTTCTACGT GCCCTGTGC ATCATGGCCT TCGTGTACCT GCGGGTGTTC CGCGAGGCC 840
AGAAGCAGGT GAAGAAGATC GACAGTGC GAGCGCGTTT CCTCGCGCGC CCAGCGCGGC 900
CGCCCTCGCC CTCGCCCTCG CCCGTCCCGG CGCCCGCGCC GCGCGCCGGA CCCCCGCGCC 960
CCGCGCGCGC CGCCGCCACC GCCCGCTGG CCAACGGGCG TCGGGGTAAG CGGCGGCCCT 1020
CGCGCCTCGT GGCCTACGC GAGCAGAAG CGCTCAAGAC GCTGGGCATC ATCATGGCGG 1080
TCTTTCACGT CTGCTGGCTG CCCTTCTTCC TGGCCAACGT GGTGAAGGCC TTCCACCGCG 1140
AGCTGGTGCC CGACCGCCTC TTCGTCTTCT TCAACTGGCT GGGCTACGCC AACTCGGCCT 1200
TCAACCCCAT CATCTACTGC CGAGCCCGCG ACTTCCGCAA GGCCTTCCAG GGACTGCTCT 1260
GCTGCGCGCG CGCGCTGCG AC CGCGCGCC AC CGAGACCCA CGGAGACCG CGCGCGCCT 1320
CGGGCTGTCT GCGCGCGGCC GACCCCGCG CATCGCCCGG GCGCGCCTCG GACGACGAG 1380
ACGACGATGT CGTCGGGGCC AC CGCGCGCG CGCGCCTGCT GAGAGCCTGG GCGCGCTGCA 1440
ACGCGCGGGC GCGCGCGGAC AGCGACTCGA GCCTGGACGA GCGGTGCCCG CCGCGCTTGG 1500
CCTCGGAATC CAAGGTGTAG GCGCGCGCG GGGCGCGCGA CTCGGGCGAC GGCTTCCAG 1560
GGGAACGAGG AGATCTGTGT TTACTTAAGA CCGATAGCAG GTGAACCTGA AGCCACAAT 1620
CCTCGTCTGA ATCATCGAG GCAAAGAGAA AAGCCACGGA CCGTTGCACA AAAAGGAAAG 1680
TTTGGGAAGG GATGGGAGAG TGGCTTGTG ATGTTCTTGG TTG

```

30

Protein Accession #:

SEQ ID NO:187 PAV1 Protein sequence

AA011176

35

40

```

1 11 21 31 41 51
MGAGVLVLGA SEPNLSSAA PLPDGAATAA RLLVPASPPA SLPPASESP EPLSQWTAG 60
MGLLMALIVL LIVAGNVLVI VAIKTPRLQ TLTNLFIMSL ASADLVMLL VVFPFATIVV 120
WGRWEYGSFF CELWTSVDVL CVTASITELC VIALDRYLAI TSPFRYQSL TRARARGLVC 180
TVWAISALVS FLPILMWWR AESDEARRCY NDPKCCDFVT NRAYATASSV VSFYVPLCIM 240
AFVYLRVFRE AQQVKKIDS CERRFLGGPA RPPSPSPSPV PAPAPPPGPP RPAAAAATAP 300
LANGRAGKRR PSRLVALREQ KALKTLGIM GVFTLCWLPF FLANVVKAFH RELVPDRLFV 360
FFNWLGYANS AFNPITYCRS PDKRAKQGL LCCARRAARR RHATHGDRPR ASGCLARPGP 420
PPSPGAASDD DDDVVVGATP PARLLEPWAG CNGGAADSD SSLDEPCRPG FASESKV

```

45

SEQ ID NO:188 BCO2 DNA sequence

Nucleic Acid Accession #:

AJ400877

Coding sequence:

81-3080 (underlined sequences correspond to start and stop codons)

50

55

60

65

70

75

80

```

1 11 21 31 41 51
GGCGTCCGCG CACACCTCCC CGCGCGCGCG CCGCCACCGC CCGCACTCCG CCGCCTCTGC 60
CCGCAACCGC TGAGCCATCC ATGGGGGTG CGGGCCGCAA CCGTCCCGGG GCGGCTGGG 120
CGGTGCTGCT GCTGCTGCTG CTGCTGCCGC CACTGCTGCT GCTGGCGGGG GCCGTCCCGC 180
CGGGTCGGGG CCGTGCCGCG GGGCCGAGG AGGATGTAGA TGAGTGTGCC CAAGGGCTAG 240
ATGACTGCCA TGCCGACGCG CTGTGTCAGA ACACACCCAC CTCCTACAAG TGCTCCTGCA 300
AGCCTGGGTA CCAAGGGGAA GGCAGGCAGT GTGAGGACAT CGATGAATGT GGAAATGAGC 360
TCAATGGAGG CTGTGTCCAT GACTGTTTGA ATATTCCAGG CAATTATCGT TGAATTGTT 420
TTGATGGCTT CATGTGGCTT CATGACGGCT ATAATTGTCT TGATGTGGAC GAGTGCCTGG 480
AGAACAATGG CCGCTGCCAG CATACTGTG TCAACGTAT GGGGAGCTAT GAGTGTGCT 540
GCAAGGAGGG GTTTTCTCTG AGTGACAATC AGCACACCTG CATTACCGC TCGGAAGAGG 600
GCCTGAGCTG CATGAATAAG GATCACGGCT GTAGTCACAT CTGCAAGGAG GCCCAAGGG 660
GCAGCGTCGC CTGTGAGTGC AGGCCTGGTT TTGAGCTGGC CAAGAACCAG AGAGACTGCA 720
TCITGACCTG TAACCATGGG AACGGTGGGT GCCAGCACTC CTGTGACGAT ACAGCCGATG 780
GCCAGAGTG CAGTGCCAT CCACAGTACA AGATGCACAC AGATGGGAGG AGCTGCCTTG 840
AGCGAGAGGA CACTGTCCTG GAGGTGACAG AGAGCAACAC CACATCAGTG GTGGATGGGG 900
ATAAACGGGT GAAACGGCGG CTGCTCATGG AAACGTGTGC TGTCAACAAT GGAGGCTGTG 960
ACCGACCTG TAAGGATACT TCGACAGGTG TCCACTGCAG TTGCTCTGTT GGATCACTC 1020
TCCAGTTGGA TGGGAAGACA TGTAAGATA TTGATGAGTG CCAGACCCGC AATGGAGGTT 1080
GTGATCATTT CTGCAAAAAC ATCGTGGGCA GTTTTGACTG CGGCTGCAAG AAAGGATTTA 1140
AATTATTAAC AGATGAGAA GCTTGCCAA AGATGGATGA GTGCTCTTTG GATAGGACCT 1200
GTGACCACAG CTGCATCAAC CACCCTGGCA CATTGTCTG TGCTTGAAC CGAGGGTACA 1260
CCCTGTATGG CTTCACCCAC TGTGGAGACA CCAATGAGTG CAGCATCAAC AACGGAGGCT 1320
GTCAGCAGGT CTGTGTGAAC ACAGTGGGCA GCTATGAATG CCAGTGCCAC CCTGGGTACA 1380
AGCTCCACTG GAATAAAAAA GACTGTGTGG AAGTGAAGGG GCTCCTGCC ACAAGTGTGT 1440
CACCCTGTG TCCCTTGCAC TCGGTAAGA GTGGTGGAGG AGACGGGTGC TTCCTCAGAT 1500
GTCACCTTGG CATTACCTC TCTTCAGATG TCACCACCAT CAGGACAAGT GTAACCTTTA 1560
AGCTAAATGA AGGCAAGTGT AGTTTGAAAA ATGCTGAGCT GTTTCGCCAG GGTCTGCGAC 1620
CAGCACTACC AGAGAAGCAC AGCTCAGTAA AAGAGAGCTT CCGCTACGTA AACCTTACAT 1680
GCAGCTCTGG CAAGCAAGTC CCAGGAGCCC CTGCCCAGC AAGCACCCT AAGGAAATGT 1740
TTATCACTGT TGAGTTTGG CTTGAACTA ACCAAAAGGA GGTGACAGCT TCTTGTGACC 1800

```

TGAGCTGCAT CGTAAAGCGA ACCGAGAAGC GGCTCCGTAA AGCCATCCGC ACGCTCAGAA 1860
 AGGCCGCTCCA CAGGGAGCAG TTTCACCTCC AGCTCTCAGG CATGAACCTC GACGTGGCTA 1920
 AAAAGCCTCC CAGAACATCT GAACGCCAGG CAGAGTCTCT TGGAGTGGGC CAGGGTCATG 1980
 CAGAAAAACCA ATGTGTCTAGT TGCAGGGCTG GGACCTATTA TGATGGAGCA CGAGAACGCT 2040
 GCATTTTATG TCCAAATGGA ACCTTCCAAA ATGAGGAAGG ACAAATGACT TGTGAACCAT 2100
 GCCCAAGACC AGGAAATTCT GGGGCCCTGA AGACCCCAAG AGCTTGAAT ATGTCTGAAT 2160
 GTGGAGGTCT GTGTCAACCT GGTGAATATT CTGCAGATGG CTTTGCACCT TGCCAGCTCT 2220
 GTGCCCTGGG CAGGTTCCAG CCTGAAGCTG GTCGAACCTC CTGCTTCCCT TGTGGAGGAG 2280
 GCCTTGCCAC CAAACATCAG GGAGCTACTT CCTTTCAGGA CTGTGAAACC AGAGTTCAAT 2340
 GTTCACCTGG ACATTTCTAC AACACCACCA CTCACCGATG TATTCGTTCG CCAGTGGGAA 2400
 CATACCAGCC TGAATTGGA AAAAAATAAT GTGTTTCTTG CCCAGGAAAT ACTACGACTG 2460
 ACTTGTATGG CTCACAAAC ATAACCCAGT GTAAAAACAG AAGATGTGA GGGGAGCTGG 2520
 GAGATTTCAC TGGGTACATT GAATCCCCAA ACTACCCAGG CAATTACCCA GCCAACACCG 2580
 AGTGTACGTG GACCATCAAC CCACCCCCCA AGCGCCGCAT CCTGATCGTG GTCCTGAGA 2640
 TCTTCTGCC CATAGAGGAC GACTGTGGGG ACTATCTGGT GATGCGGAAA ACCTCTTCAT 2700
 CCAATTCTGT GACAACATAT GAAACCTGCC AGACCTACGA ACGCCCATC GCCTTCACCT 2760
 CCAGGTCAAA GAAGCTGTGG ATTCAAGTCA AGTCCAATGA AGGGAACAGC GCTAGAGGGT 2820
 TCCAGTCCC ATACGTGACA TATGATGAGG ACTACCAGGA ACTCATTGAA GACATAGTTC 2880
 GAGATGGCAG GCTCTATGCA CTGAGAACC ATCAGGAAAT ACTTAAGGAT AAGAACTTA 2940
 TCAAGGCTCT GTTGTATGTC CTGCCCCATC CCCAGAACTA TTTCAAGTAC ACAGCCCAGG 3000
 AGTCCCGAGA GATGTTTCCA AGATCGTICA TCCGATTGCT ACGTTCCAAA GTGTCCAGGT 3060
 TTTTGAGACC TTACAAATGA CTCAGCCAC GTGCCACTCA ATACAAATGT TCTGCTATAG 3120
 GGTGTGGTGG ACAGAGCTGT CTTCTTCTG CATGTCAGCA CAGTCGGGTA TTGCTGCCTC 3180
 CCGTATCAGT GACTCATAG AGTTCAATTT TTATAGATAA TACAGATATT TTGGTAAATT 3240
 GAACTTGGTT TTTCTTCCC AGCATCGTGG ATGTAGACTG AGAATGGCTT TGAGTGGCAT 3300
 CAGCTTCTCA CTGCTGTGGG CGGATGTCTT GGATAGATCA CGGGCTGGCT GAGCTGGACT 3360
 TTGGTCAGCC TAGGTGAGAC TCACCTGTCC TTCTGGGGTC TTAACCTCC TCAAGGAGTC 3420
 TGTAGTGGAA AGGAGGCCAC AGAATAAGCT GCTTATCTG AAACCTCAGC TTCTCTAGC 3480
 CCGGCCCTCT CTAAGGGAGC CCTCTGCACT CGTGTGACAG CTCTGACCAG GCAGAACAGG 3540
 CAAGAGGGA GGAAGAGGAGA CCCCTGCAAG CTCCTCCAC CCACCTTGA ACCTGGGAGG 3600
 ACTCAGTTTC TCCACAGCCT TCTCCAGCCT GTGTGATACA AGTTTGATCC CAGGAACCTG 3660
 AGTTCTAAGC AGTGCTCGTG AAAAAAAAAA GCAGAAAGAA TTAGAAATAA ATAAAACTA 3720
 AGCACTCTG GAGACAT

SEQ ID NO:189 BCO2 Protein sequence
 Protein Accession #: CAB92285

1 11 21 31 41 51
 | | | | |
 MGVAGRNRPG AAWAVLLLLL LLPPLLLLAG AVPPGRGRGAA GPQEDVDECA QGLDDCHADA 60
 LCQNTPTS YK CCKPGYQGE GRQCEDIDEC GNELNGGCVH DCLNIPGNYR CTCFDGFMLA 120
 HDGHNCILDVD ECLENNGGCQ HTCNVNMGSY ECCCKEGFFL SDNQHTCIHR SEEGLSCMNK 180
 DHGCSHICKE APRGSVACEC RPFELAKNQ RDCILTCHNG NGGQCHSCDD TADGPECSCH 240
 PQYKMHDTDPV SCLEREDTVL EVTESNTTSV VDGDKRVRKR LLMETCAVNN GGCDRTCKDT 300
 STGVHSCSPV GFTLQDGGT CKDIDECQTR NGGCDHFCKN IVGSFDCGCK KGFKLLTDEK 360
 SCQDVDECSL RDTCDHSCIN HPGTFACACN RGYTLYGFTH CGDTNECSIN NGGQCVQCVN 420
 TVGSYECQCH PGYKLHWNKK DCVEVKGLLP TSVSPRVSLH CGKSGGGDGC FLRCHSGIHL 480
 SSDVTIIRTS VTFKLNEGKC SLKNAELFPE GLRPALEPKH SSVKESFRYV NLTCSSGKQV 540
 PGAPGRPSTP KEMFITVEFE LETNQKEVTA SCDLSCIVKR TEKRLRKAIR TLRKAVHREQ 600
 FHLQLSGMNL DVAKKPPRTS ERQAESCGVG QGHAENQCVS CRAGTYDGA RERCILCPNG 660
 TFONEGQMT CEPKPRPGNS GALKTPAWN MSECGLCQP GEYSADGFAP CQLCALGTFQ 720
 PEAGRTSCFP CGGGLATKHQ GATSFQDCET RVQCSPGHFY NTTHRCIRC PVGTYPQPEF 780
 KNNVSCPGN TTIDFDGSTN ITQCKNRRCG GELGDFGTGY ESPNYPGNYP ANTECTWTIN 840
 PPKRRILIV VPEIFLPIED DCGDYLVMRK TSSNSVTTY ETCQTYERPI AFTSRSKLW 900
 IQFKNSENGS ARGFQVPYVT YDEDYQELIE DIVRDGRLYA SENHQEILKD KKLKALFDV 960
 LAHPQNYFKY TAQESREMFP RSFIRLLRSK VSRFLRPYK

SEQ ID NO:190 BFG1 DNA sequence
 Nucleic Acid Accession #: AF007170
 Coding sequence: 1-1725 (underlined sequences correspond to stop codon)

1 11 21 31 41 51
 | | | | |
 AAGGAGGCGG CCTCCGGGAA AAGCGACCGC AGGACTCTG AGAGCAGCCT CCATGAGGCC 60
 CTGGACCAGT GCATGACCGC CCTGGACCTC TTCCTACCA ACCAGTTCTC AGAAGCACTC 120
 AGCTACCTCA AGCCAGAAAC CAAGGAAAGC ATGTACCACT CACTGACATA TGCCACCATC 180
 CTGGAGATGC AGGCCATGAT GACCTTGTAC CCTCAGGACA TCCTGCTTGC CGGCAACATG 240
 ATGAAGGAGG CACAGATGCT GTGTCAAGG CACCGGAGGA AGTCTTCTGT AACAGATTCC 300
 TTCAGCAGCT TGGTGAACCG CCCACGCTG GGCCAATTCA CTGAAGAAGA AATCCACGCT 360
 GAGGTCTGCT ATGCAGAGTG CTGTCTGCAG CGAGCAGCCC TGACCTTCTT GCAGGACGAG 420
 AACATGGTGA GCTTCATCAA AGGCGGCATC AAAGTTCGAA ACAGCTACCA GACCTACAAG 480
 GAGCTGGACA GCCTTGTTCA GTCCTCACA TACTGCAAGG GTGAGAACCA CCCGCACITT 540
 GAAGGAGGAG TGAAGCTTGG TGTAGGGGCC TTCAACCTGA CACTGTCCAT GCTTCCTACT 600
 AGGATCTGTA GGTCTTGTGA GTTGTGGGG TTTTCAGGAA ACAAGGACTA TGGGCTGCTG 660
 CAGCTGGAGG AGGGAGCGTC AGGGCACAGC TTCCGCTCTG TGCTCTGTGT CATGCTCTG 720
 CTGTGCTACC ACACCTTCTT CACCTTCGTG CTCGGTACTG GGAACGTCAA CATCGAGGAG 780
 GCCGAGAAGC TCTTGAAGCC CTACCTGAAC CGGTACCCTA AGGGTGCCAT CTTCTGTTC 840
 TTTGACGGA GGAATGAAGT CATTAAAGGC AACATTGATG CAGCCATCCG GCGTTTCGAG 900

GAGTGTCTGT AGGCCAGCA GCACTGGAAG CAGTTTCACC ACATGTGCTA CTGGGAGCTG 960
 ATGTGGTGCT TCACCTACAA GGGCCAGTGG AAGATGTCTT ACTTCTACGC CGACCTGCTC 1020
 AGCAAGGAGA ACTGCTGGTC CAAGGCCACC TACATTTACA TGAAGGCCGC CTACCTCAGC 1080
 ATGTTTGGGA AGGAGGACCA CAAGCCGTTT GGGGACGACG AAGTGGAAAT ATTTGAGCT 1140
 GTGCCAGGCC TGAAGCTCAA GATTGCTGGG AAATCTCTAC CCACAGAGAA GTTTGCCATC 1200
 CGGAAGTCCC GGCCTACTT CTCCTCCAAC CCTATCTCGC TGCCAGTGCC TGCTCTGAA 1260
 ATGATGTACA TCTGGAACGG CTACGCCGTG ATTGGGAAGC AGCCGAAACT CACGGATGGG 1320
 ATACTTGAGA TTATCACTAA GGCTGAAGAG ATGCTGGAGA AAGGCCCAGA GAACGAGTAC 1380
 TCAGTGGATG ACGAGTGCTT GGTGAAATTG TTGAAAGGCC TGTGTCTGAA ATACCTGGGC 1440
 CGTGTCAGG AGGCCGAGGA GAATTTTAGG AGCATCTCTG CCAATGAAAA GAAGATTAAA 1500
 TATGACCACT ACTTGATCCC AAACGCCCTG CTGGAGCTGG CCTGCTGCTT TATGGAGCAA 1560
 GACAGAAACG AAGAGGCCAT CAAACTTTTG GAATCTGCCA AGCAAACTA CAAGAATTAC 1620
 TCCATGGAGT CAAGGACACA CTTTGAATC CAGGCAGCCA CACTCCAAGC CAAGTCTTCC 1680
 CTAGAGAACA GCAGCAGATC CATGGTCTCA TCAGTGTCTT TGTAGCTTTG TGCAGCAGTT 1740
 CCGGGCTGGA AGACAGAGAC AGCTGGACAG AGCTCTGAA AACATTTCAA AATACCCCT 1800
 CCCCCTGCCC TGCCCTGCCT TTGGGGTCCA CCGGCACTCC AGTTGGATGG CACAACATAG 1860
 TGTATCCGTG CAGAAGCCGA GCTGGCATTT TCACCAAGTGT AGCCAAGGGC CTTTGCCAAG 1920
 GGCAGAGCAG GTGGAGCCCT CTGCCTGCCC TATCACACAT ACGGGTACTT GCTTTTCACT 1980
 GTGATGTTTA AGAGAATGTA TGAACAGTTT ACATTTTCCT TAGAAATACA TTGATGGGAT 2040
 CACAGTTGGC TTAAAAACC AACACAATC AACCACTGT AAGTCTTTGT CTTACCTAT 2100
 TATCATCTGG AGGTAAATCT CTTATATGA TGATGCCAAA GGCAAAATG CTTTCAAAT 2160
 TCAGCAAGTT CTCAGCTTGT GTGACGGAAG GTCCTCAGA GGACCTGAGG AATGCCTGGG 2220
 AGAGGCTAAG CCTCAGGCTT CAATGCTTCT GGGGTGGGC ATGAGGATGT ACACAGACAC 2280
 CCATACCTT ACTACTACA CTTCAATTCA CTCCTTTTGT AAATTTCCAA TTTAAAAATC 2340
 AAGCAGCTCT TTTTAGTGAG ATAAATCTG AGCTCTTCTG TAGAAAAATC AATCTCTACC 2400
 AGTAGAAAAAT GCCAGGGCTT GATGGAAGAG CTGTGTAGCC CTTTCTATGC CAAAGCCAGG 2460
 AAATTTGGGG GGCAGGAGGA GGTCTCAGA ATCCAGTCTG TATCTTTGCT GTATGCCAAA 2520
 CTGAAACCAC TGGGAATAAT TTATGAAACA TAAAAATCTT CTGTACTTCA CTCCAAGGTA 2580
 CATTGCTTA CTGACAGCAT TTTTGTTAAA ACTGTTATTC TTGAAAAAAA AAAAAAAAAA 2640
 AA

SEQ ID NO:191 BFG1 Protein sequence

Protein Accession #: AAC39582

1 11 21 31 41 51
 | | | | |
 MTALDLFLTN QFSEALSYLK PRTKESMYHS LTYATILEMQ AMMTFDPQDI LLAGNMMKEA 60
 QMLCQRHRRK SSVTDSFSSL VNRPTLGQFT EEEIHAECY AECLLQRAAL TFLQDENMVS 120
 FIKGGIKVRN SYQTYKELDS LVQSSQYCKG ENHPHFEGGV KLGVGAFNLT LMLPTRLR 180
 LLEFVGSGN KDYGLLQLEB GASGHSFRSV LCVMLLLCYH TFLTFVLGTG NVNIEEAEL 240
 LKPYLNRYPK GAIFLFFAGR IEVIGKNIDA AIRFEECCE AQQHWKQFHH MCVWELMWCF 300
 TYKQGWKMSY FYADLLSKEN CWSKATYIYM KAAVLSMFGK EDHKPFGDDE VELFRAVPGL 360
 KLKIAGKSLP TEKFAIKSR RYFSSNPISL PVPALMMYI WNGYAVIGKQ PKLTDGILEI 420
 ITKAEEMLEK GPENEYSVDD ECLVKLLKGL CLKYLGRVQE AEENFRSISA NEKKIKYDHY 480
 LIPNALLELA LLLMEQDRNE EAIKLLESAK QNYKNYSMES RTHFRIQAAAT LQAKSSLENS 540
 SRSMVSSVSL

SEQ ID NO:192 BFO6 DNA sequence

Nucleic Acid Accession #: NM_032583

Coding sequence: 1-4044 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 ATGACTAGGA AGAGGACATA CTGGGTGCC AACTCTTCTG GTGGCTCGT GAATCGTGGC 60
 ATCGACATAG GCGATGACAT GGTTCAGGA CTTATTATA AAACCTATAC TCTCCAAGAT 120
 GGCCCTGGA GTCAGCAAGA GAGAAATCCT GAGGCTCCAG GGAGGGCAGC TGTCCACCG 180
 TGGGGGAAGT ATGATGTGTC CTTGAGAACC ATGATTCCCT TCCGTCCCAA GCCGAGGTTT 240
 CTTGCCCCCC AGCCCTGGA CAATGCTGGC CTGTTCTCCT ACCTACCGT GTCATGGCTC 300
 ACCCGCTCA TGATCCAAAG CTTACGGAGT CGCTTAGATG AGAACACCAT CCCTCCACTG 360
 TCAGTCCATG ATGCCTCAGA CAAAAATGTC CAAAGGCTTC ACCGCTTTG GGAAGAAGAA 420
 GTCTCAAGGC GAGGGAATTGA AAAAGCTTCA GTGCTTCTGG TGATGCTGAG GTTCCAGAGA 480
 ACAAGGTGTA TTTTCGATG ACTCTGGGC ATCTGCTTCT GCATTGCCAG TGTACTCGGG 540
 CCAATATTGA TTATACCAAA GATCTGGAA TATTCAGAAG AGCAGTTGGG GAATGTTGTC 600
 CATGGAGTGG GACTCTGCTT TGCCCTTTT CTCTCCGAAT GTGTGAAGTC TCTGAGTTTC 660
 TCCTCCAGTT GGATCATCAA CCAACGCACA GCCATCAGGT TCCGAGCAGC TGTTTCTCTC 720
 TTTGCCCTTG AGAAGCTCAT CCAATTAAAG TCTGTAATAC ACATCACCTC AGGAGAGGCC 780
 ATCAGCTTCT TCACCGGTGA TGTAACCTAC CTGTTTGAAG GGGGTGTGCTA TGGACCCCTA 840
 GTACTGATCA CCTGGCATC GCTGGTATC TGCAGCATT CTTCCTACTT CATTATTGGA 900
 TACACTGCAT TTATGCCAT CTTATGCTAT CTCTGTTT TCCCACTGGC GGTATTCATG 960
 ACAAGAATGG CTGTGAAGGC TCAGCATCAC ACATCTGAGG TCAGCGACCA GCGCATCCGT 1020
 GTGACCAGTG AAGTTCTCAC TTGCATTAAG CTGATTAATA TGTACACATG GGAGAAACCA 1080
 TTGCAAAAA TCAATTGAAGG TATGAAAGT CTGACTTTCT GCTCCAAACC TGGTGATGGC 1140
 ATGGCCTTCA GCATGCTGGC CTCTTGAAT CTCTTCCGGC TGTCAGTGTT CTTTGTGCCT 1200
 TTGCAAGTCA AAGTCTCAC GAATTCCAAG TCTGCAGTGA TGAGGTTCAA GAAGTTTTC 1260
 CTCCAGGAGA GCCCTGTTT CTATGTCCAG ACATTACAAG ACCCCAGCAA AGCTCTGGTC 1320
 TTTGAGGAGG CCACCTTGTG ATGGCAACAG ACCTGTCCCG GGATCGTCAA TGGGGCACTG 1380
 GAGCTGGAGA GGAACGGGCA TGCTTCTGAG GGGATGACCA GGCCTAGAGA TGCCCTCGGG 1440

5 CCAGAGGAAG AAGGGAACAG CCTGGGCCCA GAGTTGCACA AGATCAACCT GGTGGTGTCC 1500
 AAGGGGATGA TGTTAGGGGT CTGCGGCAAC ACGGGGAGTG GTAAGAGCAG CCTGTTGTCA 1560
 GCCATCCTGG AGGAGATGCA CTGTCTCGAG GGCTCGGTGG GGTGTCAGGG AAGCCTGGCC 1620
 10 TATGTCCCCC AGCAGGCTGT GATCGTCAGC GGGAAACATCA GGGAGAACAT CCTCATGGGA 1680
 GGCGCATATG ACAAGGCCCG ATACCTCCAG GTGTCCCACT GCTGCTCCCT GAATCGGGAC 1740
 CTGGAACCTT TGCCCTTTGG AGACATGACA GAGATTGGAG AGCGGGGCCT CAACCTCTCT 1800
 GGGGGGACAGA AACAGAGGAT CAGCCTGGCC CGCGCCGTCT ATTCCGACCG TCAGATCTAC 1860
 CTGCTGGACG ACCCCTGTCT TGCTGTGGAC GCCCAGCTGG GGAAGCACAT TTTTGAGGAG 1920
 15 TGCAATTAAGA AGACACTCAG GGGGAAGACG GTCGTCTGG TGACCCACCA GCTGCAGTAC 1980
 TTAGAATTTT GTGGCCAGAT CATTTTGTG GAAAAATGGGA AAATCTGTGA AAATGGAAGT 2040
 CACAGTGAGT TAATGCAGAA AAAGGGGAAA TATGCCCAAC TTATCCAGAA GATGCACAAG 2100
 GAAGCCACTT CGGACATGTT GCAGGACACA GCAAAGATAG CAGAGAAGCC AAAGGTAGAA 2160
 AGTCAGGCTC TGGCCACCTC CCTGGAAGAG TCTCTCAACG GAAATGCTGT GCCGGAGCAT 2220
 CAGCTCACAC AGGAGGAGGA GATGGAAGAA GGCTCCTTGA GTTGGAGGGT CTACCACCAC 2280
 20 TACATCCAGG CAGTGGAGG TTACATGGTC TCTTGATATA TTTCTTCTT CGTGGTGGTG 2340
 ATCGTCTTCT TAACGATCTT CAGCTTCTGG TGGCTGAGCT ACTGGTTGGA GCAGGGCTCG 2400
 GGGACCAATA GCAGCCGAGA GAGCAATGGA ACCATGGCAG ACCTGGGCAA CATTGCAGAC 2460
 AATCCTCAAC TGTCCTTCTA CCAGCTGGTG TACGGGCTCA ACGCCCTGCT CCTCATCTGT 2520
 GTGGGGGTCT GCTCCTCAGG GATTTTCACC AAAGTCACGA GGAAGGCATC CACGGCCCTG 2580
 25 CACAACAAGC TCTTCAACAA GGTTTTCCGC TGCCCATGTA GTTCTTTGA CACCATCCCA 2640
 ATAGGCCGGC TTTTGAACCT CTTCGACAGG GACTTGGAA AGCTGGACCA GCTCTTGCCC 2700
 ATCTTTTCAG AGCAGTTCCT GGTCTGTCC TTAATGGTGA TCGCCGTCTT GTTGATTGTC 2760
 AGTGTGCTGT TCCATATAT CCGTTAATG GGAGCCATAA TCATGGTTAT TTGCTTCATT 2820
 TATTATATGA TTTTCAAGAA GGCCATCGGT GTGTTCAAGA GACTGGAGAA CTATAGCCGG 2880
 30 TCTCTTTAT TCTCCACAT CTTCAATTCT CTGCAAGGCC TGAGTCCAT CCATGTCTAT 2940
 GGAAAACTG AAGACTTCAT CAGCCAGTTT AAGAGGCTGA CTGATGCGCA GAATAACTAC 3000
 CTGCTGTGT TCTATCTTC CACACGATGG ATGGCATTGA GGCTGGAGAT CATGACCAAC 3060
 CTGTGACCT TGGCTGTGC CCGTGTCTG GCTTTTGGCA TTCTCTCCAC CCCCTACTCC 3120
 TTAAAGTCA TGGCTGTCAA CATGTGCTG CAGCTGGCGT CCAGCTTCCA GGCCACTGCC 3180
 CGGATTGGCT TGGAGACAGA GGCACAGTTC ACGGCTGTAG AGAGGATACT GCAGTACATG 3240
 35 AAGATGTGT TCTCGGAAGC TCCTTTACAC ATGGAAGGCA CAAGTTGTCC CCAGGGGTGG 3300
 CCACGACATG GGGAAATCAT ATTTAGGAT TATCACATGA AATACAGAGA CAACACACCC 3360
 ACCGTGCTC ACGGCATCA CTTGACCATC CGCGGCCACG AAGTGGTGGG CATCGTGGGA 3420
 AGGACGGGCT CTGGGAAGTC CTCCTGGGC ATGGCTCTCT TCCGCCTGGT GGAGCCCATG 3480
 GCAGGCCGGA TTTCTATTGA CGGCGTGGAC ATTTGCAGCA TCGGCTGGA GGAAGTGGG 3540
 40 TCCAAGCTCT CAGTGTATCC TCAAGATCCA GTGCTGCTCT CAGGAACCAT CAGATTCAAC 3600
 CTAGATCCCT TTGACCGTCA CACTGACCAG CAGATCTGGG ATGCCTTGA GAGGACATTC 3660
 CTGACCAAGG CCATCTCAAA GTTCCCCAAA AAGTGCATA CAGATGTGGT GGAACACGGT 3720
 GGAACATCT CTGTGGGGGA GAGGCAGCTG CTCTGCATTG CCAGGGCTGT GCTTCGCAAC 3780
 TCCAAGATCA TCTTATCGA TGAAGCCACA GCCTCCATG ACATGGAGAC AGACACCCCTG 3840
 45 ATCCAGCGCA CAATCGTGA AGCCTTCAG GGCTGCACCG TGCTCGTCAT TGCCACCGT 3900
 GTCACCACTG TGCTGAAGTG TGACCACATC CTGGTTATGG GCAATGGGAA GGTGGTAGAA 3960
 TTTGATCGG CGGAGGTAT GCGGAAGAAG CTTGGGTCTT TGTTGCGAG CCTCATGGCC 4020
 ACAGCACTT CTTCATGAG ATAAGGAGAT GTGGAGACT CATGAGGCT GGCAGCTGAG 4080
 CTCAGAGGTT CACACAGGTG CAGCTTCGAG GCCACAGTC TGCGACCTTC TTGTTTGGAG 4140
 50 ATGAGAACTT CTCCTGGAAG CAGGGGTAAA TGTAGGGGGG GTGGGGATTG CTGGATGGAA 4200
 ACCCTGGAAT TCTGCTATTG ATGGCTCTCA AGACCTTAGA ACCCCAGAAC CATCTAAGAC 4260
 ATGGGATCA GTGATCATGT GGTTCCTCT TTAACCTACA TGCTGAATAA TTTTATAATA 4320
 AGGTAAGAG TTAATGTTT CTGATCTGTG TTAGAAGTGY TGCAAATGCT GTACTGACTT 4380
 TGTAATAAT AAAACTAAGG AAAACTCAAA AAAAAAAAAA AAAAAA

SEQ ID NO:193 BFO6 Protein sequence

Protein Accession #: NP_115972.1

55 1 11 21 31 41 51
 MTRKRTYWVP NSSGGLVNRG IDIGDDMVSG LIYKTYTLQD GPWSQQERNP EAPGRAAVPP 60
 WGKYDAALRT MIPFRPKPRF PAPQPLDNAG LFSYLTVSWL TPLMIQSLRS RLDENTIPPL 120
 60 SVHDASDKNV QRLHRLWEEV VSRRIEIKAS VLLVMLRFQR TRLIFDALLG ICFIASVLG 180
 PILIIPKILE YSEEQLGNVV HGVGLCFALF LSECYKSLSF SSSWIINQRT AIRFRAAVSS 240
 FAFEKLIQFK SVIHITSGEA ISFFTGDVNY LFEGV CYGPL VLITCASLVI CSISSYFIIG 300
 YTAFAILCY LLVFLAVFVM TRMAVKAQHH TSEVSDQRI RVTSEVLTCIK LIKMYTWEKP 360
 65 FAKIEGMEIS LTFCKPGDGM MAFSMLASLN LLRLSVFFVP IAVKGLTNSK SAVMRFKKFF 420
 LQESPVFYVQ TLQDPSKALV FEEATLSWQQ TCPGVNGAL ELERNHASE GMTRPRDALG 480
 PEEEGNSLGP ELHKINLVVS GKMMLGVCGN TSGVKSSLLS AILEEMHLE GSVGVQGSLS 540
 YVPQAWIVS GNIRENILMG GAYDKARYLQ VHHCCSLNRD LELLFPDGMT EIGERGLNLS 600
 GGQKQRISLA RAVYSRQIY LLDDPLSAVD AHVGKHIFEE CIKTLRGT VVLVTHQLQY 660
 LEFCGQILL ENGKICENTG HSELMOQKKGK YAQLIQKMHK EATSDMLQDT AKIAEKPVE 720
 70 SQALATSLLE SLNGNAVPEH QLTQEEMEE GSLSWRVYHH YIQAGGYMV SCIIFFVVL 780
 IVFLTIFSW WLSYWLQGS GTNSRESNG TMADLGNLAD NPQLSFYQLV YGLNALLIC 840
 VGVCSGIFT KVTBKASTAL HNKLFNKVF CPMSFFDTIP IGRLLNCFAG DLEQLDQLLP 900
 IFSEQLVLS LMVIAVLLIV SVLSPYILLM GAIMVICFI YMMFMKKAIG VFKRLNYSR 960
 SPLFSLHLS LQGLSIHVV GKTEFISQF KRLTDAQNNY LLLFLSSTRW MALRLEIMTN 1020
 75 LVTLAVALFV AFGISSHPY FKVMVAVNIV LCLASSQATA RIGLETEAQF TAVERILQYM 1080
 KMCVSEAPLH MEGTSCPGGW PQHGEIHFQD YHMKYRDNT TVLHGINLTI RGHEVVGVIVG 1140
 RTGSGKSLG MAFRLVPEM AGRILIDGVD ICSIGLEDLR SKLSVIPQDP VLLSGTIRFN 1200
 LDPFDRHTDQ QIWDALERTF LTKAISKFPK KLHTDVVENG GNFSVGERQL LCIARAVLRN 1260
 SKILIDEAT ASIDMETDTL IQRTIREAFQ GCTVLVIAHR VTVLNCDDHI LVMGNKGKVV 1320
 FDRPEVLRKK PGSFLAALMA TATSSLR

SEQ ID NO:194 BHB8 DNA sequence

Nucleic Acid Accession #:

AA983251

Coding sequence:

1-1749 (underlined sequences correspond to start and stop codons)

5	1	11	21	31	41	51	
	ATGCTGTCTG	GCTTCTTGAT	GAGTCCCAGT	ACCCAGCACA	GAGCACAGTA	CACTCCCCGA	60
10	GGAAAGAAAC	TTCCGTGGGA	GGCTTCCATC	GGTGCGCACA	CCTCCCCGAGG	GCGAGGCAGC	120
	GACCGGGAGA	GGGAGAGCCG	GCCGGAGGCT	GCCGGGCTCC	TGTGGGACCG	CGCTGCAGCC	180
	GGGAGAGCGG	AGAAGGGGAA	CCGGGGCGAG	CCGCCCGCCT	GGATCCGCGC	CCAGCAGCAG	240
	CCGCGGCGCG	CGCCAGCTGG	GCAGGCTCCC	GGGACTGCGG	CTGGGGGCGC	GCAGGACCCT	300
15	CGCCTGCGTC	CTGGACGTTT	CCGGGGGAGG	GTCCGGTTGC	CAGTGAACCC	TCCAGAGGCT	360
	TCCGGACGAC	AGCCCCGGGG	GCCTTCTGAC	TGCATCCCCA	GATTTCCATC	AGCGAGTGCA	420
	ACTCATAAGG	CAGTCCCATA	GGGACCCGGG	CCACCGGCTG	AGGACGGGGA	TGGCTTAGGA	480
	GCTCCTGGAC	CTAGGGCCCC	GCCTCGTCGC	CTCCTGGGCG	TCGCGGCAGA	GGGAGGTGGC	540
	CCGCGCGGAA	AGCCCGCGCG	GACAGTCACT	GACGAGGCCC	GGGGGTGCGC	GGGGCCACGA	600
20	CTTCTCGGAG	ACCGTCTGCG	GCTCTCTGGA	GACGCGCTGT	CCGCGCCCAG	GGTGGTGCCA	660
	TGTGGGGCGC	TCGCGCGCTC	TCCGTCTCCT	CATCCTGGAA	CGCCGCTTCG	CTCCTGCAGC	720
	TGCTGCTGGC	TGCGCTGCTG	GCGCGGGGGG	CGAGGGCCCA	GCGGCGAGTA	CTGCCACGGC	780
	TGGCTGGACG	CGCAGGGCGT	CTGGCGCATC	GGCTTCCAGT	GTCCCGAGCG	CTTCGACGGC	840
	GGCGACGCCA	CCATCTGCTG	CGGCAGCTGC	GCCTTGCCTG	ACTGCTGCTC	CAGCGCCGAG	900
25	GCGCGCGTGG	ACCAGGGCGG	CTCGCACAAT	GACCGCCAGC	AGGGCGCTGG	CGAGCCTGGC	960
	CGGGCGGACA	AAGACGGGGC	CCGACGGCTC	GGCAGGGCTT	CATGTCTTAG	GGGTACCCAA	1020
	GGAGACGGCG	AGGTTGCGCC	CCACCCCGTG	AGGGCCTGGC	AGCGGTGCTC	CCCTGAAGGC	1080
	TCCCCGAAAG	GAAGGCAGCT	CCTCAGGGCT	TTCCCGGGGC	TGCTGCCCCG	TGCCAGACGC	1140
	CGCGGATTCC	CATCTTCTCC	ACGCGGCGGC	CCCTCTCCCC	TGCAGCGGCC	CGCCTTGCCC	1200
30	ATCTACGTGC	CCTTCCCTCAT	TGTTGGCTCC	GTGTTTGTGC	CCTTTATCAT	CTTGGGGTCC	1260
	CTGTTGGCAG	CGTCTTCTGT	CAGATGTCTC	CGGCCTAAGC	AGGATCCCCA	GCAGAGCCGA	1320
	GCCCCAGGGG	GTAACCGCTT	GATGGAGACC	ATCCCCATGA	TCCCCAGTGC	CAGCACCTCC	1380
	CGGGGGTTCG	CCTCACGCCA	GTCCAGCACA	GCTGCCAGTT	CCAGCTCCAG	CGCCAACTCC	1440
	GGGGCCCCGG	CGCCCCCAAC	AAGGTCACAG	ACCAACTGTT	GCTTGCCGGA	AGGGACCATG	1500
35	AACAACGTGT	ATGTCAACAT	GCCACGAAAT	TTCTCTGTGC	TGAACGTGTA	GCAGGCCACC	1560
	CAGATTGTGC	CACATCAAGG	GCAGTATCTG	CATCCCCCAT	ACGTGGGGTA	CACGGTGCAG	1620
	CAGCACTGCT	TGCCCCATGC	AGCTGTGCCA	CCTTTTCATG	ACGGCCTGCA	GCCTGGCTAC	1680
	AGGCAGATTG	AGTCCCCCTT	CCCTCACACC	AACAGTGAAC	AGAAGATGTA	CCCAGCGGTG	1740
40	ACTGTATAAC	CGAGAGTCAC	TGTTGGGTTT	CTTTACTGAA	GGGAGACGAA	GGCAGGGGTG	1800
	GATTCTCGAG	GTGGAAGTCC	GCACATGTGC	GTGGTATTTA	TGGCACGATT	CCTTTGGATG	1860
	GCTTCATTGT	CCCCCAGACT	GTATGAAAAC	ATCTCCGAAT	TAGCATTTCT	GGATATGTTT	1920
	ATCCAGGGT	ATCATTGATT	TATGATGGAA	AACCGGCCCT	AGCTGGAGAT	GACTGTGATG	1980
	TGCTGTATGG	GTGTATAACA	AATGCTTGAG	TCCGAAGTGC	CCTTGAGATA	TGGTTGACGA	2040
45	AAGAATTTTA	TAAACTGATA	AATTAAGGAT	TTTTATATAT	TGTTTATTAT	TATTTCTTTT	2100
	TTGTTGTTGA	CTGCACAGGA	TCAAAATGCC	TGTTATCTCC	CTTTTACTGG	GACTTTTTTT	2160
	TTTTTTTTTT	TTTTTTTTTAA	TCAGACAGGG	TCTTGCTCTG	TTGCCCAGGC	TGGAGTGCAG	2220
	TGGTGCATC	TCGGCTCACT	GCAACTFCAG	CCTCCTGGAT	TCAGGCAACA	CTCCTGCCCT	2280
	AGCCTCCCAC	GTGGCTGGGA	TTACAGGTGC	CTGCCCCCAT	GGCTAATTTT	TTGTATTTTT	2340
50	TGTAGAGATG	GGGTTTCACC	ATGTTGGCTG	GGCTGGTCTC	ACTCTCCTGA	CCTCAAGCAA	2400
	TCTGCCCTGC	TCAGCCTCCC	AAAGTGCTGG	GATTACAGGC	GTGAGCCACC	GCCCCCAGCC	2460
	TGAGCCTTTT	TTTTTTTCTA	ATGCATCCAA	GGTTAAGGGG	AAGACGCAAA	TAACAGGACT	2520
	ATTCTAAAAG	GAAACCTGTT	TGAACCTCTG	GAGATCAGTC	ATCAGTCTCA	GTATTCCACA	2580
	GGCACACCTT	AATTTTCATT	TAAAAAGATA	TATATATTTT	GTCTATTTTT	GTGCTTTTTG	2640
55	GGGCCTATTT	TGTGCTTTTT	TACCTTATGT	AGAGATCTTA	TTACAAAGTG	ATTTTCTACA	2700
	TTAAAAAGAG	ACTGAAATAA	ATTGTATAGT	TACTTAACTA	ATGAAGACAT	TTCAGAACTC	2760
	TGGGATGATT	TTAATCTTGA	AGTAGTAGGT	GGTATAGTCA	TAAAACCATT	CATCCCCCTC	2820
	TTGATTGTAT	CTTAATTTTC	TGGCTTTAAG	GTGACATCTG	AGAGGTAATG	CATTCTTTTT	2880
	TATATTGAAA	TCATAAACTA	TCACCCGCTG	CTTCTCTGAG	TTACTTTTAA	TTTTGCCTTG	2940
	TGGTTATGGT	TTGGCGTTTC	CTTCTGTTTG	GTPTTCAGAG	CCCCATGTCT	ATATAGTCTC	3000
60	GAGTGCAAGT	AATTACTATA	CTTGTAATG	AAGATCAGTA	TTTCTGCCCT	GATCTGATAA	3060
	AAAAATTTTC	TTGTCTTAGT	TATAAAAATT	CAAAGAAATG	TGTTACAAAG	ATACTTAGTA	3120
	TAGCTCCTCA	GCCATAACCT	GAGACTTGGG	ATGAAATTTA	AACCAGATAC	GATTTACTTT	3180
	CGAGATCATA	AGGCTTTTTA	TACTCTTGTT	ATCAAAATGG	CTTATTTTTT	AGGCACTAAG	3240
	GATTGTTAAG	AGAAAAGCTT	TTCAACGAAG	GATTGCCTTT	CTTCTCCAC	ACTGTTCTTG	3300
65	ATTTCCTCTC	TCPTTCAGGC	CTCAACAGGC	ACTGTATTCA	TTGCCAATGT	TCCAAATTAT	3360
	CAAAATTCAG	TGAATTTATT	TGTGTGTTCT	TTACTTATAT	AAAAAAGAT	AACTTTAAGG	3420
	ATGTGCAAGT	ACATTTCCAA	CTGCTAGCAC	AACCAGTATT	TTGTAATTAA	ACAAATCGCT	3480
	GATGGGTATG	GTCTTCTACA	CATTTATGTC	TATAGATATC	TATCGATCAT	CTTCTATATC	3540
	TGTTTCTATG	CTGAATAATG	TAAACCAGT	GTTGGCAATT	GGTATCATCA	ATGATACTCA	3600
70	TTTTTTAATA	ACCAAAAGCA	GGGAAAAATC	ATTTTACTTA	TTAATAAATA	TTTTATGATG	3660
	TGAAAAAATA	AAAAAATAAA	AAAAAATAAA				

SEQ ID NO:195 BHB8 Protein sequence

Protein Accession #: none found

5
10
15
20

1	11	21	31	41	51	
MLSGFLMSPS	TQHRAQYTPG	GKKLPWEAST	GAHTSRGRGS	DRERESRPEA	AGLLWDRAAA	60
GEAEKGNRGE	PPAWIRAQQQ	PREPPAGQAP	GTAAGGAQDF	RLRPGRSRGR	VRLPVKPPEA	120
SGRQPRGPSD	CIPRFPSASA	THKAVPKGTG	PPAEDGDGLG	APGPRARRRR	LLGVAAEGSG	180
PRGKRRTVTS	DEARGSFPGPR	LLGDRPALSG	DALSAPRVVF	CGALAARPSF	HPGTPLRSCS	240
CCWLRCWRRG	RGPSGEYCHG	WLDAGQVWRI	GFQCPERFDG	GDATICCGSC	ALRYCCSSAE	300
ARLDQGGCDN	DRQQGAGEPG	RADKDGPRRL	GRASCLRGTQ	GDGEGAPPFV	RAWQRCSPEG	360
SPKGRQLLRA	FPGLLEPRARR	RGFFSSPRGG	PSPLQRPALF	IYVPFLIVGS	VVFVFIILGS	420
LVAACCCRCL	RPKQDPQQR	APGGNRLMET	IPMIPSASTS	RGSSSRQSST	AASSSSSANS	480
GARAPFTRSQ	TNCCLEPEGTM	NNVYVNMPTN	FSVLNCQQAT	QIVPHQGQYL	HPPVVGYTVQ	540
HDSVPMTAVF	PFMDGLQPGY	RQIQSPFPHT	NSEQKMPAV	TV		

SEQ ID NO:196 CQA5 DNA SEQUENCENucleic Acid Accession #: AA088458
Coding sequence:

862-1995 (underlined sequences correspond to start and stop codons)

25
30
35
40
45
50
55
60
65
70
75

1	11	21	31	41	51	
GCCCTTGGAC	ACTGACATGG	ACTGAAGGAG	TAGAATGGAG	CACGAGGACA	CTGACATGGA	60
CTGAAGAAAA	AGGAGCTGGA	GCAGGAGAAG	GAGGTGCTGC	TGCAGGGTTT	GGAGATGATG	120
GGCGGGGGCC	GCGACTGGTA	CCAGCAGCAG	CTGCAACGAG	TGCAGGAGCG	CCAGCGCCGC	180
CTGGGCCAGA	GCAGAGCCAG	CGCCGACTTT	GGGGCTGCAG	GGAGCCCCCG	CCCACTGGGG	240
CGGCTACTGC	CCAAGGTACA	AGAGGTGGCC	CGGTGCCTGG	GGGAGCTGCT	GGCTGCAGCC	300
TGTGCCAGCC	GGGCCCTGCC	CCCGTCTCTC	TCCGGGCCCC	CCTGCCCTGC	CCTGACGTCC	360
ACCTCACCCC	CGGTCTGGCA	GCAGCAGACC	ATCCTCATGC	TGAAGGAGCA	GAACCGACTC	420
CTCACCCAGG	AGGTGACCGA	GAAGAGTGAG	CGCATCACGC	AGCTGGAGCA	GGAGAAGTCG	480
GGCTCATTTA	AGCAGCTGTT	TGAGGCCCGC	GCCCTGAGCC	AGCAGGACGG	GGGACCTCTG	540
GATTCCACCT	TCATCTAGTC	CTTGTGGGCC	GCCTGGGGCC	CCAGGGCCAG	CCTGGCACTC	600
AGCCCTTCGA	GGGTGGGCGC	CCCATTCGCAC	CCACCCTCTC	TGGCTGGAGA	CCCCCGGCAG	660
GCCAGGCGAC	AGTCCCGGAG	TGGGCGCCTT	CCTGCCGCCC	TTGCCAGATG	GGCTCCCCAG	720
GGCTGCCCCC	GGCTGGTCCC	CGCACCGAGC	GCTTGACTCC	GTTTGGGCTC	CTGGTTGYTG	780
ACATGGGCTG	GGGGCTCTCT	TGAGTCCGCA	TAGTCCGCAG	CTACTACTGG	CCGCTGTCTAG	840
TGGACAGTGG	GGTACCCCTC	CATGAGTTAG	CGTCCCCCGG	TTTCCAGCGG	TGCGCGCCTG	900
GGTCCCATCT	TCAGGGAAGG	GCATGCCCCA	CGCCAGGCTG	CACCTTCCAA	AACGGGCAGC	960
AGAGGGCGCG	GGGCGGCTCC	GACGCGGGTC	CAAGGGCAGC	TTCCCGCTCA	ACCAGGGCAC	1020
CAGGACGAGG	TGGCTGTAGC	TCGGACGGAC	GGAAGTAGAT	GGAGGGGGTG	GGGACGGCCT	1080
GTAAGCGGGG	GGTGCCTGCC	TGGCTGGGGA	GCCCCAGGGA	TAGCGGTTCG	ACTTCAGGTT	1140
CTGGCCCAAG	CTGAGGGACC	CTGGCTGCAG	CGGATCGGCA	CGCCGGGTGG	GCGAGAGCTT	1200
GGCCTGCATG	TGCCCTCCAC	AGACCCTGGG	GTGATGGCCT	TCCCCCTCTT	GGCCGGGACG	1260
TTGCCCCACG	TTGAGTCCCA	CACAACATCC	TGTGAGCCTG	GCTCCCCAGG	AGGGCCCCCA	1320
GACAGCTCCC	AGGCACGTCA	TAGGCAAAAG	CTGTTTCCCC	CGACTCAGGA	TTTCCAAGGC	1380
CTGGGGTCTC	GCTCACCCCC	CTTTGCTCTC	ACGCCCAGCC	TGTCCCCAGG	TTTCAGCTGG	1440
GAGAGGCCAC	CTCCCTCAGC	CAAGGAAAAC	GAGAACCCCC	AGGGTACAGG	AGGAGGCTGG	1500
GGCAGGTCCC	CTTGGGTGTC	ACTCCCTCAG	CCCCTGCCCA	GGCCCACTCC	CGCTGGTGCT	1560
GGAGTACGCA	CTGGTGGGGG	GGCCCTGCTC	AGCCCAACCT	GGAGGGTCCC	AGTGTACCCA	1620
GAACAGGGGG	CACGGCAACA	GCATCGATGG	GTTCTGCAGC	CCAGGGCCCC	CGATGCGGGG	1680
TCAGTGTGTG	TGGGGCGCAG	GGCCTCCGAT	CGGGGGTCAG	TGCGTGGGGG	GCGCAGGGCC	1740
CCCGATGCGG	GGTCAGTGCG	TGGGGGGCGC	AGGGCCCCCT	CGTGTCCAGG	GCACTTTGGT	1800
ACACTGTCCC	ACAAGGCACC	TGTCTCAGAG	GAGGGGCCCT	GGCAGGCAGC	GTGGCAACTC	1860
CCTTCCGGAG	CCCAGCTCCA	TGCTAACCCT	CCCACAGCAA	CCCCACAGAG	CCACATTTCC	1920
TGCTGCACCT	GGTCTGCAGG	GGTGTCCCG	GACAGGCCCA	AGTCAGCCCA	GCATGCAGCT	1980
GCCCTCTTAC	CCTGAAGATG	GGAGTGGGCT	TTCCAGGGGA	CATAAGGATG	TCAGGCCCTG	2040
ACCTCTCTGG	CAGGAAGGGG	TGCAGGTCTT	GAGGGCCTGT	GCCCCACAGC	CCCAAGCACC	2100
AGGTGGACTG	CAGCGCAGTG	GGTGGGCCAG	TGGCAGCCAG	GGAGAAGCCC	CCCCGTACGA	2160
GGCTGGGGTC	TGCCCACCCAG	GGCCTCCCCA	CGTCTGCCCT	TGAGGGTGCC	TGCCATGCCC	2220
TGGGGGATCC	TGGCATCTTT	ACTGGACTGG	AAGCAGGAGA	CAGAACAGTG	TCTGTCCCGG	2280
GGTGACTTCA	TCAGGAGACC	GGCCACATAG	AGCTGGACCC	CGCAGCTGAA	GCGGAAATGT	2340
GAGACAGGCT	GGCACCTCCG	GAAAAACTGC	CTTTCAGCCT	TGGTCTTCCG	TGCAAGGTGA	2400
AAAGAAATAG	GTCCCTCCAG	TTTACAGCTT	GAAATCAGGC	TAGTGAAGTG	CCCTGGAGAC	2460
CACGAGGGGA	GAATTPTAAAG	GCCCCGGCTG	GCAGGGTCTA	GGTGGCTGGC	AGAGGCACAT	2520
GCAGACCTTG	CCTGGAGCCT	GCCCTAGGAC	GCTGGGCGGG	TCAGTCTCCG	TGCAGGATGT	2580
GAGCAGCGTC	CCTGGGCTCT	ATCCGCGAGG	TGCCAGTAGC	GTGTGCAGGT	ACATACACGT	2640
GGGTGCACAC	TGTGATGACA	CCCCGAAATG	TCTCAGGATG	TTGAAATGTG	TCCTTGGGGG	2700
CAGAAGTGTC	CCCAAGTTAG	AATCTGCCCC	AGAGGAACAC	ATCCACACCA	GGCCTCAGGA	2760
TTTGTGTGTT	ATCAAGTTCC	AAGGAAAAGG	AACATCTCAG	CCGGGCGTGG	TGGTTACAGC	2820
CTGGAATCCC	AGCACTTGAG	GCCAGGAGTT	CCAGAGCAGC	CTGGGCAACG	CAGTGAGAGA	2880
CCCCATCTCT	ACAARAAAAA	AAAAAGAAAG	AAAGAAAATG	AGAGATCCAG	GTTTAAAAAT	2940
TCATAAACAC	CACAAGGAAA	CAATACACTA	TGAGACCCAG	CAGAAGCAAC	AGATTGACTC	3000
TAGACCCAGA	TACTAGAAAT	ATCAGAGAGA	ATATAAAGTA	ACAGTGTTTT	ATATATCTAA	3060
AGAAATAAAA	GAGATTTCTG	GAAACATGAA	AAAAAA			

SEQ ID NO:197 LBG2 DNA SEQUENCE

Nucleic Acid Accession #: X63629
 Coding sequence: 54-2543 (start and stop codons are underlined)

```

5      1      11      21      31      41      51
      |      |      |      |      |      |
GCGGAACACC GGCCCCCGGT CGCGGCAGCT GCTTCACCCC TCTCTCTGCA GCCATGGGGC 60
TCCCTCGTGG ACCTCTCGCG TCTCTCTCC TTCTCCAGGT TTGCTGGCTG CAGTGGCGGG 120
CCTCCGAGCC GTGCCGGGCG GTCTTCAGGG AGGCTGAAGT GACCTTGGAG GCGGGAGGGC 180
10    CGGAGCAGGA GCCCGGCCAG GCGCTGGGGA AAGTATTTCAT GGGCTGCCCT GGGCAAGAGC 240
CAGCTCTGTT TAGCACTGAT AATGATGACT TCACTGTGCG GAATGGCGAG ACAGTCCAGG 300
AAAGAAGGTC ACTGAAGGAA AGGAATCCAT TGAAGATCTT CCCATCCAAA CGTATCTTAC 360
GAAGACACAA GAGAGATTGG GTGGTTGCTC CAATATCTGT CCCTGAAAAA GGCAAGGGTC 420
CCTTCCCCCA GAGACTGAAT CAGCTCAAGT CTAATAAAGA TAGAGACACC AAGATTTTCT 480
15    ACAGCATCAC GGGGCCGGGG GCAGACAGCC CCCCTGAGGG TGTCTTCGCT GTAGAGAAGG 540
AGACAGGCTG GTTGTGTGTT AATAAGCCAC TGGACCGGGA GGAGATTGCC AAGTATGAGC 600
TCTTTGGCCA CGTGTGTGCA GAGAATGGTG CCTCAGTGGA GGACCCCATG AACATCTCCA 660
TCATCGTGAC CGACCAGAAT GACCACAAGC CCAAGTTTAC CCAGGACACC TTCCGAGGGA 720
GTGCTTAGA GGGAGTCTTA CCAGGTACTT CTGTGATGCA GGTGACAGCC ACAGATGAGG 780
20    ATGATGCCAT CTACACCTAC AATGGGGTGG TTGCTTACTC CATCCATAGC CAAGAACCAA 840
AGGACCCACA CGACCTCATG TTCACAATTC ACCGGAGCAC AGGCACCATC AGCGTCATCT 900
CCAGTGGCCT GGACCGGGAA AAGTCCCTG AGTACACACT GACCATCCAG GCCACAGACA 960
TGGATGGGGA CGGTCCACC ACCACGGCAG TGGCAGTAGT GGAGATCCTT GATGCCAATG 1020
ACAATGCTCC CATGTTTGGC CCCCAGAAAT ACGAGGCCCA TGTGCCTGAG AATGCAAGTG 1080
25    GCGATGAGGT GCAGAGGCTG ACGGTCACTG ATCTGGACGC CCCCACCTCA CCAGCGTGGC 1140
GTGCCACCTA CTTATCATG GCGCGTGACG ACGGGGACCA TTTTACCATC ACCACCCACC 1200
CTGAGAGCAA CCAGGGGATC CTGACAACCA GGAAGGGTTT GGATTTTGAG GCCAAAAACC 1260
AGCACACCTT GTACGTTGAA GTGACCAACG AGGCCCTTTT TGTGCTGAAG CTCCCAACCT 1320
CCACAGCCAC CATAGTGGTC CACGTGGAGG ATGTGAATGA GGCACCTGTG TTTGTCCAC 1380
30    CCTCCAAAGT CGTTGAGGTC CAGGAGGGCA TCCCCTGAGT GGAGCCTGTG TGTGCTACA 1440
CTGCAAGAAG CCCTGACAAG GAGAATCAAA AGATCAGCTA CCGCATCCTG AGAGACCCAG 1500
CAGGGTGGCT AGCCATGGAC CCAGACAGTG GGCAGGTGAC AGCTGTGGGC ACCCTCGACC 1560
GTGAGGATGA GCAGTTTGTG AGGAACAACA TCTATGAAGT CATGGTCTTG GCCATGGACA 1620
35    ATGGAAGCCC TCCCACCACT GGCACGGGAA CCCTTCTGCT AACACTGATT GATGTCAACG 1680
ACCATGGCCC AGTCCTGAG CCCCCTGAGA TCACCATCTG CAACCAAAGC CTTGTGCGCC 1740
ACGTGCTGAA CATCAGGAC AAGGACCTGT CTCCCACAC CTCCCCTTTC CAGGCCACG 1800
TCACAGATGA CTCAGACATC TACTGGACGG CAGAGGTCAA CGAGGAAGGT GACACAGTGG 1860
TCTTGTCCCT GAAGAAGTTC CTGAAGCAGG ATACATATGA CGTGACCTT TCTCTGTCTG 1920
40    ACCATGGCAA CAAAGAGCAG CTGACGGTGA TCAGGGCCAC TGTGTGCGAC TGCCATGGCC 1980
ATGTCGAAAC CTGCCCTGGA CCTGGAAAG GAGGTTTCAT CCTCCCTGTG CTGGGGGCTG 2040
TCTGGCTCT GCTGTCTCTC CTGCTGTGTC TGCTTTTGTG GGTGAGAAAG AAGCGGAAGA 2100
TCAAGGAGCC CCTCTACTC CCAGAAGATG ACACCCGTGA CAACGTCTTC TACTATGGCG 2160
AAGAGGGGGG TGCGAAGAG GACCAGGACT ATGACATCAC CCAGTCCAC CGAGGTCTGG 2220
45    AGGCCAGGCC GTCCGAGCGG GCGTCCCTGA GCTCCCTCAC CTCCTCCGCC TCCGACCAAG 2460
ACCAAGATTA CGATTATCTG AACGAGTGGG GCAGCCGCTT CAAGAAGCTG GCAGACATGT 2520
ACGGTGGCGG GGAGGACGAC TAGCGGCCT GCCTGCAGGG CTGGGGACCA AACGTCAAGC 2580
50    CACAGAGCAT CTCAAAGGGG TCTCAGTTC CCCTTCAGCT GAGGACTTCG GAGCTTGTC 2640
GGAAGTGGCC GTAGCAACTT GGCAGGAGCA GGCTATGAGT CTGACGTTAG AGTGGTTGCT 2700
TCCTTAGCCT TTACGATGG AGGAATGTGG GCAGTTTGAC TTCAGCACTG AAAACCTCTC 2760
CACTGGGCC AGGGTTGCCT CAGAGGCCAA GTTTCAGAA GCCTCTTACC TGCCGTAAAA 2820
55    TGCTCAACCC TGTGTCTGG GCCTGGGCCT GCTGTGACTG ACCTACAGTG GACTTTCTCT 2880
CTGGAATGGA ACCTTCTTGT GCCTCTGTGT GCAACTTAAT TTTTITTTT AATGCTATCT 2940
TCAAAACGTT AGAGAAAGTT CTCAAAAGT GCAGCCCAAG GCTGCTGGGC CCAGTGGCCG 3000
TCTGCAATTT CTGTTTCCA GACCCCAATG CCTCCATTG GGATGGATCT CTGCGTTTTT 3060
60    ATACTGAGTG TGCCTAGGTT GCCCTTATT TTTATTTTC CCGTGTGCGT TGCTATAGAT 3120
GAAGGGTGAG GACAATCGTG TATATGTACT AGAATTTTT TATTAAAGAA A

```

SEQ ID NO:198 LBG2 Protein sequence:
 Protein Accession #: CAA45177

```

65      1      11      21      31      41      51
      |      |      |      |      |      |
MGLPRGPLAS LLLLQVCWLQ CAASEPCRAV FREAEVTLEA GGAEQEPGQA LGKVFMGCPG 60
QEPALFSTDN DDFTVRNGET VQERRSLKER NPLKIFPSKR ILRRHKRDWV VAPISVPENG 120
70    KGPFPQRLNQ LKSNKDRDRT IFYSITGPGA DSPPEGVFAV EKETGWLLLN KPLDREEIAK 180
YELFGHAYSE NGASVEDPMN ISIVTDQND HKPKFTQDTF RGSVLEGVLP GTSVMQVTAT 240
DEDDAIYTYN GVVAYSIHSQ EPKDPHDLMF TIHRSTGTIS VISSGLDREK VPEYTLTIQA 300
TDMGDGDSST TVAVVVEILD ANDNAPMFDQ QKYEAHVPEV AVGHEVQRLT VTDLDAPNSP 360
75    AWRATYLMIG GDDGDHFTIT THPESNQIL TTRKGLDFA KNOHTLYVEV TNEAPFVLKL 420
PTSTATIVVH VEDVNEAPVF VPFSKVVEVQ EGIPTGEPVC VYTAEDPDKE NQKISYRILR 480
DPAGWLAMDP DSGQVTA VGT LDREDEQFVR NNIYEVMLA MDNGSPPTTG TGTLLTLID 540
VNDHGPVPEP RQITCNQSP VRHVLNITDK DLSPHTSFPQ AQLTDDSDIY WTAEVNEEGD 600
TVVLSLKKFL KQDTYDVHLS LSDHGNKEQL TVIRATVDCD HGHVETCPG WKGGFILPVL 660
GAVLALLFLL LVLLLLVRKK RKIKEPLLP EDDTRDNVfy YGEEGGGEED QDYDITQLHR 720

```

GLEARPEVVL RNDVAPTIIP TPMYRPRPAN PDEIGNFIE NLKAANTDPT APPYDILLVF 780
DYESGSDAA SLSSLTSSAS DQDQDYDYLN EWGSRFKKLA DMYGGGEDD

SEQ ID NO:199 OBI5 DNA SEQUENCE

5 Nucleic Acid Accession #: NM_012152
Coding sequence: 43-1104 (underlined sequences correspond to start and stop codons)

```

10 1      11      21      31      41      51
    |      |      |      |      |      |
    CTCTTTTAAA TTTCTTTCTA GGATGTTTAC TTCTTCTCCA CAATGAATGA GTGTCACTAT 60
    GACAAGCACAT TGGACTTTTTT TTATAATAGG AGCAACACTG ATACTGTGCGA TGACTGGACA 120
    GGAACAAAGC TTGTGATTGT TTTGTGTGTT GGGACGTTTT TCTGCCTGTT TATTTTTTTT 180
    TCTAATTCCTC TGGTCATCGC GGCAGTGATC AAAAACAGAA AATTTCATTT CCCCTTCTAC 240
    TACCTGTTGG CTAATTTAGC TGCTGCCGAT TTCTTCGCTG GAATTGCCTA TGTATTCTCTG 300
    ATGTTTAAACA CAGGCCCACT TTCAAAAACCT TTGACTGTCA ACCGCTGGTT TCTCCGTCAG 360
    GGGCTTCTGG ACAGTAGCTT GACTGCTTCC CTCACCAACT TGCTGGTTAT CGCCGTGGAG 420
    AGGCACATGT CAATCATGAG GATGCGGGTC CATAGCAACC TGACCAAAAA GAGGGTGACA 480
    CTGCTCATTT TGCTTGTCTG GGCCATCGCC ATTTTATATG GGGCGGTCCC CACACTGGGC 540
    TGGAATTGCC TCTGCAACAT CTCTGCCCTGC TCTTCCCTGG CCCCATTATA CAGCAGGAGT 600
    TACCTTGTCT TCTGGACAGT GTCCAACCTC ATGGCCTTCC TCATCATGGT TGTGGTGTAC 660
    CTGCGGATCT ACGTGTACGT CAAGAGGAAA ACCAAGTCTT TGCTCTCCGA TACAAGTGGG 720
    TCCATCAGCC GCCGGAGGAC ACCCATGAAG CTAATGAAGA CCGTGTATGAC TGTCTTAGGG 780
    GCGTTTGTGG TATGCTGGAC CCCGGGCCCTG GTGGTTCTGC TCCTCGACGG CCTGAAGTGC 840
    AGGCAGTGTG GCGTGCAGCA TGTGAAAAGG TGGTTCTCTG TGCTGGCGCT GCTCAACTCC 900
    GTCGTGAACC CCATCATCTA CTCTTACAAG GACGAGGACA TGTATGGCAC CATGAAGAAG 960
    ATGATCTGCT GCTTCTCTCA GGAGAACCCA GAGAGGCGTC CCTCTCGCAT CCCCTCCACA 1020
    GTCCTCAGCA GGAGTGACAC AGGCAGCCAG TACATAGAGG ATAGTATTAG CCAAGGTGCA 1080
    GTCTGCAATA AAAGCACATC CTAAACTCTG GATGCCTCTC GGGCCACCCA GGTGATGACT 1140
    GTCTTAGG
  
```

SEQ ID NO:200 OBI5 Protein sequence:

Protein Accession #: NP_036284

```

35 1      11      21      31      41      51
    |      |      |      |      |      |
    MNECHYDKHM DFFYNRSNTD TVDDWTGTKL VIVLCVGTFF CLFIFFSNLS VIAAVIKNRK 60
    PHFPFYLLA NLAAADFFAG IAYVPLMFNT GPVSKTLTVN RWFLRQGLLD SSLTASLTNL 120
    LVIAVERHMS IMRMVHNSL TKKRVTLLIL LVWAIAIFMG AVPTLGNWCL CNISACSSLA 180
    PIYSRSYLVF WTVSNLMAFL IMVVVYLRIV VYVKRKTNVL SPHTSGSISR RRTFMKLMKT 240
    VMTVLGAFVW CWTPLGLVLL LDGLNCRQCG VQHVKRWFLL LALLNSVVPN IIYSYKDEDM 300
    YGTMMKMICC FSQENPERRP SRIPSTVLSR SDTGSQYIED SISQGVNCK STS
  
```

SEQ ID NO:201 PAA6 DNA SEQUENCE

Nucleic Acid Accession #: AA569531
Coding sequence: 1-504 (underlined sequences correspond to start and stop codons)

```

50 1      11      21      31      41      51
    |      |      |      |      |      |
    ATGACCTACA GTTACTCATT TTTCAGGCCT GAGTTGATCG TTAATCATCT TAATTATGTT 60
    CATTCTGAAG CCAACAGGAG AACCAAGACC AAAACTTTAT TGTCTCTGCT TTCATTCTCTT 120
    GATGAAACCT CTGGACTAAG CACACATCTT CCTTGTATTAT CTCTCTCAAA GGAGTGTGGA 180
    GTGCTTCAATC TGGACATCCA CGGGAAGAAG GAAGACATGA GAATCACCCA ACAGTCTPTCC 240
    CAGCTATACC TGTGGGACAT GGGTGGTTTT ACAATATTTA AGAACCTGTG GATGAGCCTC 300
    ATACCAGAG GGAACAAACG CTCCCCAAAA AGAGTTACAG AAACCATCCT GAGAGATTTT 360
    AAGCAGAAGC AAAGTTCAAA GATCCAAGAG GAGAGACGAA GAGAGTCTGC AGGACCAAAC 420
    CTCTCTTCAT TCTGGTTTGT GGGGAATGCT GGAAGAGGAG ACAGGCCCCA GATTGTGGCA 480
    GGAAGTAAAC AGTTTTCAGG CTGAGGCCAA TCTGAGCAGG AACATTCCAA TATTCTTTCA 540
    GCTACGTTGT CCCAGCACAT CACTGGTTAA CCTTTTATGT CCACCATTTG TGGATTTTCA 600
    AGCTACTTGT CAATGGTGAA TATTGATCAT CATCATTATC TACTGAGCTG CTACCATATC 660
    CCAGCTACTC CTTGATGTTT GTTCATTATT TTCTCAACAC TCAGCATATT TGCAATATGT 720
    TATGTAATAT CACAGACAAG GAAACTGAAC GCAGAAATGT TTTATTTCTT GCCAAACATC 780
    ACATAGAGAT GAACATGAA ACCGATTTGA AACCAGGATT GTCTGATTCC AACATCTCTG 840
    GGTCTCTTTT CACTCTGATA TGCTGCAATT AAAAAGCCAT TTCTAAGACT GT
  
```

SEQ ID NO:202 PAA6 Protein sequence:

Protein Accession #: none found

```

70 1      11      21      31      41      51
    |      |      |      |      |      |
    MTYSYSFFRP ELIVNHLNVV HSEANRRTKT KTLLSLLSFL DETSGLSTHL PCLSLSKECG 60
    VLHLDIHGK EDMRITQQSS QLYLWDMGGF TIFKNLWMSL IPRGNKRSPK RVTETILRDF 120
    KQKQSSKIQE ERRRESAGPN LSSFVFWGNA GRGDRPQIWA GSKQFSG
  
```

SEQ ID NO:203 PAB2 DNA SEQUENCE

Nucleic Acid Accession #:
Coding sequence:

XM_050197

310-1971 (underlined sequences correspond to start and stop codons)

5 1 11 21 31 41 51
| | | | | |
TCACACGTGC CAAGGGGCTG GCTCAGCGGA ACCAGCCTGC ACGCGCTGGC TCCGGGTGAC 60
AGCCGCGCGC CTCGGCCAGG ATCTGAGTGA TGAGACGTGT CCCCACTGAG GTGCCCCACA 120
GCAGCAGGTG TTGAGCATGG GCTGAGAAGC TGGACCGGCA CCAAGGGGCT GGCAGAAATG 180
GGCGCTGGC TGATTCTCTAG GCAGTTGGCG GCAGCAAGGA GGAGAGGCGC CAGCTTCTGG 240
AGCAGAGCCG AGACGAAGCA GTTCTGGAGT GCCTGAACGG CCCCCTGAGC CCTACCCGCC 300
TGGCCCACTA TGGTCCAGAG GCTGTGGGTG AGCCGCTGC TCGGCACCG GAAAGCCAG 360
CTCTTGCTGG TCAACCTGCT AACCTTTGGC CTGGAGGTGT GTTTGGCCGC AGGCATCACC 420
TATGTGCCGC CTCTGCTGCT GGAAGTGGGG GTAGAGGAGA AGTTTCATGAC CATGGTGCTG 480
GGCATTGGTC CAGTGTCTGG CCTGTCTGT GTCCCGCTCC TAGGCTCAGC CAGTGACCAC 540
TGGCGTGGAC GCTATGGCCG CCGCCGGGCC TTCTATCTGG CACTGTCTCTT GGCATCTCTG 600
CTGAGCCTCT TTCTCATCCC AAGGGCCGGC TGGCTAGCAG GGCTGCTGTG CCCGGATCCC 660
AGGCCCTTGG AGCTGGCACT GCTCATCTCT GGCCTGGGGC TGCTGGACTT CTGTGGCCAG 720
GTGTGCTTCA CTCCACTGGA GGCCCTGCTC TCTGACCTCT TCCGGGACCC GGACCACTGT 780
CGCCAGCCCT ACTCTGTCTA TGCCCTCATG ATCAGTCTTG GGGGCTGCCT GGGCTACCTC 840
CTGCCCTGCC TTGACTGGGA CACCAGTGCC CTGGCCCCCT ACCTGGGCAC CCAGGAGGAG 900
TGCCCTCTTG GCCTGCTCAC CCTCATCTTC CTCACCTGCG TAGCAGCCAC ACTGCTGGTG 960
GCTGAGGAGG CAGCGCTGGG CCCCACCGAG CCAGCAGAAG GGCTGTGCGC CCCCTCTCTG 1020
TCGCCCACT GTGTCTCATG CCGGGCCCGC TTGGCTTTCC GGAACCTGGG CGCCCTGCTT 1080
25 CCCCCTGTC ACCAGCTGTG CTGCCGATG CCCCACCC TCGCCGGCT CTTCGTGGCT 1140
GAGCTGTGCA GCTGGATGGC ACTCATGACC TTCACGCTGT TTTACACGGA TTTCTGTGGC 1200
GAGGGGCTGT ACCAGGCGCT GCCCAGAGCT GAGCCGGGCA CCGAGGCCCG GAGACACTAT 1260
GATGAAGGCG TTGCGATGGG CAGCCTGGGG CTGTTCTCTG AGTGCGCCAT CTCCTCTGGT 1320
TTCTCTCTGG TCATGGACCG GCTGGTGCAG CGATTGGGCA CTCGAGCAGT CTATTTGGCC 1380
30 AGTGTGGCAG CTTTCCCTGT GGCTGCCGGT GCCACATGCC TGTCCACAG TGTGGCCGTG 1440
GTGACAGCTT CAGCCGCCCT CACCGGGTTC ACCTTCTCAG CCTTCAGAT CCTGCCCTAC 1500
ACACTGGCCT CCCTCTACCA CCGGAGAAAG CAGGTGTTCC TGCCCAAATA CCGAGGGGAC 1560
ACTGGAGGTG CTAGCAGTGA GGACAGCCTG ATGACCAGCT TCCTGCCAGG CCTTAAGCCT 1620
35 GGAGCTCCCT TCCCTAATGG ACACGTGGGT GCTGGAGGCA GTGGCTGCT CCCACCTCCA 1680
CCCGCGCTCT GCGGGGCTTC TCCTGTGAT GTCTCCGTAC GTGTGGTGGT GGGTGAGCCC 1740
ACCGAGGCCA GGTGGTTC CCGCCGGGGC ATCTGCCTGG ACCTCGCCAT CCTGGATAGT 1800
GCCTTCTCTG GTTCTTACCA GCGCCCATCC CTGTTTATGG GCTCCATTGT CCAGCTCAGC 1860
CAGTCTGTCA CTGCCATATG GTGTCTGCC GCAGGCCTGG GTCTGGTCCG CATTTACTTT 1920
40 GCTACACAGG TAGTATTTGA CAAGAGCGAC TTGGCCAAAT ACTCAGCGTA GAAACTTCC 1980
AGCACATTGG GGTGAGGGC CTGCCCTACT GGGTCCCAGC TCCCCTGCTC TGTAGCCCC 2040
ATGGGGCTGC CCGGCTGGCC GCCAGTTTCT GTGTCTGCCA AAGTAATGTG GCTCTCTGCT 2100
GCCACCTGT GTGTCTGAGG TGCGTAGCTG CACAGCTGGG GGCTGGGGCG TCCCTCTCCT 2160
CTCTCCCGAG TCTCTAGGGC TGCCGTAGCT GAGGCCTTCC AAGGGGGTTT CAGTCTGGAC 2220
TTATACAGGG AGGCCAGAAG GGCTCCATGC ACTGGAATGC GGGGACTCTG CAGGTGGATT 2280
45 ACCCAGGCTC AGGGTTAAAC GCTAGCCCTC TAGTTGAGAC ACACCTAGAG AAGGGTTT 2340
GGGAGCTGAA TAACTCAGT CACCTGGTTT CCCATCTCTA AGCCCCCTAA CCTGCAGCTT 2400
CGTTTAAATG AGCTCTTGCA TGGGAGTTTC TAGGATGAAA CACTCCTCCA TGGGATTTGA 2460
ACATATGAAA GTTATTTGTA GGGGAAGAGT CCTGAGGGGC AACACACAAG AACAGGTCC 2520
50 CCTCAGCCCC ACAGGCACCT GTCTTTTCTG CTNGANTCCA CCCCCCCCT CTTTACCTTT 2580
TT

SEQ ID NO:204 PAB2 Protein sequence:

Protein Accession #:

XP_050197

55 1 11 21 31 41 51
| | | | | |
60 MVQRLWVSRL LRHRKAQLLL VNLLTFGLEV CLAAGITYVP PLLLEVGVEE KFM TMVLGIG 60
FVLGLVCVPL LGSASDHWRG RYGRRRPFIW ALSLGILLSL FLIPRAGWLA GLLCPPDRPL 120
ELALLILGVG LDFCGQVCF TPLEALLSDL FRDPDHCRQA YSVYAFMISL GGCLGYLLPA 180
IDWDTSALAP YLGTQBECLF GLLTLIFLTC VAATLLVAEE AALGPTEPAE GLSAPSLSPH 240
CCPCRARLAF RNLGALLPRL HQLCCRPRT LRRLFVAELC SWMALMTFTL FYTDFVGEGL 300
65 YQGVPRAPPG TEARRHYDEG VRMGSLGLFL QCAISLVFSL VMDRLVQRFQ TRAVYLASVA 360
AFPVAAGATC LSHSVAVVTA SAALTGFTFS ALQILPYTLA SLYHREKQVF LPKYRGDTGG 420
ASSEDLSMTS FLPGPKPGAP FENGHVAGG SGLLPPPPAL CGASACDVSV RVVVGEPTEA 480
RVVPGRGICL DLAILDSAFL LSQVAPSLFM GSIVQLSQSV TAYMVSAGL GLVAIYFATQ 540
VVFDKSDLAK YSA

SEQ ID NO:205 PAJ3 DNA SEQUENCE

Nucleic Acid Accession #:
Coding sequence:

AK002126

1-1593 (underlined sequences correspond to start and stop codons)

75 1 11 21 31 41 51
| | | | | |
ATGTTTCGCC GGGGGCTGCT TGCGTGGATT TCCCGGGTGG TGGTTTGTCT GGTGCTCCTC 60
TGCTGTGCTA TCTCTGTCTT GTACATGTTG GCCTGCACCC CAAAAGGTGA CGAGGAGCAG 120
80 CTGGCACTGC CCAGGGCCAA CAGCCCCACG GGGAAGGAGG GGTACCGAGC CGTCCCTCAG 180
GAGTGGGAGG AGCAGCACCG CAACTACGTG AGCAGCCTGA AGCGGCAGAT CGCACAGCTC 240

AAGGAGGAGC TGCAGGAGAG GAGTGAGCAG CTCAGGAATG GGCAGTACCA AGCCAGCGAT 300
 GCTGCTGGCC TGGGTCTGGA CAGGAGCCCC CCAGAGAAAA CCCAGGCCGA CCTCCTGGCC 360
 TTCTTGCACG CGCAGGTGGA CAAGGCAGAG GTGAATGCTG GCGTCAAGCT GGCCACAGAG 420
 TATGCAGCAG TGCCTTTTGA TAGCTTTACT CTACAGAAGG TGTACCAGCT GGAGACTGGC 480
 CTTACCCGCC ACCCCGAGGA GAAGCCTGTG AGGAAGGACA AGCGGGATGA GTTGGTGGAA 540
 GCCATTGAAT CAGCCTTGGA GACCTGAAC AATCCTGCAG AGAACAGCCC CAATCACCGT 600
 CCTTACACGG CCTCTGATTT CATAGAAGGG ATCTACCGAA CAGAAAGGGA CAAAGGAGCA 660
 TTGTATGAGC TCACCTTCAA AGGGGACCAC AAACACGAAT TCAAACGGCT CATCTTATTT 720
 CGACCATTCG GCCCCATCAT GAAAGTGAAA AATGAAAAGC TCAACATGGC CAACACGCTT 780
 ATCAATGTTA TCGTGCCTCT AGCAAAAAGG GTGGACAAGT TCCGGCAGTT CATGCAGAA 840
 TTCAGGGAGA TGTGCATTGA GCAGGATGGG AGAGTCCATC TCACTGTGTG TTACTTTGGG 900
 AAAGAAGAAA TAAATGAAGT CAAAGGAATA CTTGAAAACA CTTCCAAAGC TGCCAACTTC 960
 AGGAACTTTA CCTTCATCCA GCTGAATGGA GAATTTTCTC GGGGAAAGGG ACTTGATGTT 1020
 GGAGCCCCGT TCTGGAAGGG AAGCAACGTC CTTCTCTTTT TCTGTGATGT GGACATCTAC 1080
 TTCACATCTG AATTCCTCAA TACGTGTAGG CTGAATACAC AGCCAGGGAA GAAGGTATTT 1140
 TATCCAGTTC TTTTTCAGTCA GTACAATCCT GGCATAATAT ACGGCCACCA TGATGCAGTC 1200
 CCTCCCTTGG AACAGCAGCT GGTCAATAAG AAGGAAACTG GATTTTGGAG AGACTTTGGA 1260
 TTTGGGATGA CGTGTCACTG TCGGTACAGC TTCAATCAATA TAGGTGGGTT TGATCTGGAC 1320
 ATCAAAGGCT GGGCGGAGA GGAATGTGCAC CTTTATCGCA AGTATCTCCA CAGCAACCTC 1380
 ATAGTGGTAC GGACGCCGTG GCGAGGACTC TTCCACCTCT GGCATGAGAA GCGCTGCATG 1440
 GACGAGCTGA CCCCCGAGCA GTACAAGATG TGATGCAGT CCAAGGCCAT GAACGAGGCA 1500
 TCCACGGCC AGCTGGGCAT GCTGGTGTTC AGGCACGAGA TAGAGGCTCA CCTTCGCAAA 1560
 CAGAAACAGA AGACAAGTAG CAAAAAACA TGA

SEQ ID NO:206 PAJ3 Protein sequence:

Protein Accession #: NP_060841

1 11 21 31 41 51
 MVRRLGLAWI SRVVLLVLL CCAISVLYML ACTPKGDEEQ LALPRANSPT GKEGYQAVLQ 60
 EWEEQHRNVV SSLKRQIAQL KEELQERSEQ LRNGQYQASD AAGLGLDRSP PEKTQADLLA 120
 FLHSQVDKAE VNAGVKLATE YAAVFPDSFT LQKVYQLEGT LTRHPEEKFP RKDKRDELVE 180
 AIESALETLN NPAENSPNHR PYTASDFIEG IYRTERDKGT LYELTFKGDH KHEFKRLILF 240
 RPFGPIMKVK NEKLNMATFL INVIVPLAKR VDKFRQFMQN FREMCIEQDG RVHLTVVYFG 300
 KEEINEVKGI LENTSKAANF RNFTFIQLNG EFSRKGGLDV GARFWKGSNV LLFFCDVDIY 360
 FTSEFLNTRC LNTQPGKKVF YPVLFSQYNP GIIYGHHDV PPLEQQVLVIK KETGFWRDFG 420
 FGMTQYRSD FINIGGEDVH IKGWGEDVH LYRKYLSNL IIVRTPVRGL FHLWHEKRCM 480
 DELTPEQYKM CMQSKAMNEA SHGQLGMLVF RHEIEAHLRK QKQKTSKKK

SEQ ID NO:207 PAJ5 DNA SEQUENCE

Nucleic Acid Accession #: AF189723

Coding sequence: 1-2712 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 ATGATTCCTG TATTGACATC AAAAAAGCA AGTGAATTAC CAGTCAGTGA AGTTGCAAGC 60
 ATTCTCCAAG CTGATCTTCA GAATGGTCTA AACAAATGTG AAGTTAGTCA TAGGGCAGCC 120
 TTTTCATGGCT GGAATGAGTT TGATATTAGT GAAGATGAGC CACTGTGGAA GAAGTATATT 180
 TCTCAGTTTA AAAATCCCTT TATTATGCTG CTTCTGGCTT CTGCAGTCAT CAGTGTMTTA 240
 ATGCATCAGT TTGATGATGC CGTCAGTATC ACTGTGGCAA TACTTATCGT TGTTACAGTT 300
 GCCTTTGTTC AGGAATATCG TTCAAGAAAA TCCTTTGAAG AATTGAGTAA ACTTGTGCCA 360
 CCAGAAATGCC ATGTGTGTGCG TGAAGGAAAA TTGGAGCATA CACTTGCCCG AGACTTGGTT 420
 CCAGGTGATA CAGTTTGCTT TTCTGTTGGG GATAGAGTTC CTGCTGACTT ACGCTTGTPT 480
 GAGGCTGTGG ATCTTTCCAT TGATGAGTCC AGCTTGACAG GTGAGACAAC GCCTTGTTC 540
 AAGGTGACAG CTCTTCAGCC AGCTGCAACT AATGGAGATC TTGCATCGAG AAGTAACATT 600
 GCCTTTATGG GAACACTGGT CAGATGTGGC AAAGCAAAGG GTGTTGTCTAT TGGAAACAGGA 660
 GAAAATCTTG AATTTGGGGA GGTTTTAAAT ATGATGCAAG CAGAAGAGGC ACCAAAAACC 720
 CCTCTGCAGA AGAGCATGGA CCTCTTAGGA AAACAACCTT CTTTCTTACT CTTTGGTATA 780
 ATAGGAATCA TCATGTTGGT TGGCTGGTTA CTGGGAAAAA ATATCCTGGA AATGTTTACT 840
 ATTAGTGTAA GTTTGGCTGT AGCAGCAATT CCTGAAGGTC TCCCCATTGT GGTACAGATG 900
 ACGCTAGCTC TTGGTGTATAT GAGAATGGTG AAGAAAAGGG CCATTGTGAA AAAGCTGCCT 960
 ATTTGTTGAAA CTCTGGGCTG CTGTAATGTG ATTTGTTTCA ATAAACTGG AACACTGACG 1020
 AAGAATGAAA TGACTGTTAC TCACATATTT ACTTCAGATG GTCTGCATGC TGAGGTTACT 1080
 GGAATTGGCT ATAACTCAATT TGGGGAAGTG ATTTGTTGATG GTGATGTGTG TCATGGATTCT 1140
 TATAACCCAG CTCTTAGCAG AATTGTTGAG GCGGGCTGTG TGTGCAATGA TGCTGTAATT 1200
 AGAAACAATA CTTAAATGGG GAAGCCAACA GAAGGGCCTT TAATTGCTCT TGCAATGAAG 1260
 ATGGGCTCTG ATGGACTTCA ACAAGACTAC ATCAGAAAAA CTGAATACCC TTTTAGCTCT 1320
 GAGCAAAAGT GGAATGGCTGT TAAGTGTGTA CACCGAACAC AGCAGGACAG ACCAGAGATT 1380
 TGTTTTATGA AAGGTGCTGT CGAACAAAGTA ATTAAGTACT GTACTACATA CCAGAGCAAA 1440
 GGGCAGACCT TGACACTTAC TCAGCAGCAG AGAGATGTGT ACCAACAAAGA GAAGGCACGC 1500
 ATGGGCTCAG CGGGACTCAG AGTTCTTGTG TTGGCTTCTG GTCCTGAAC TGGACAGCTG 1560
 ACATTTCTTG GCTTGGTGGG AATCATTGAT CCACCTAGAA CTGGTGTGAA AGAAGCTGTT 1620
 ACAACACTCA TTGCCTCAGG AGTATCAATA AAAATGATTA CTGGAGATTC ACAGGAGACT 1680
 GCAAGTTGCAA TCGCCAGTCG TCTGGGATTG TATTCCAAAA CTCCCAAGTC AGTCTCAGGA 1740
 GAAGAAATAG ATGCAATGGA TGPTCAGCAG CTTTCACAAA TAGTACCAAA GGTTCAGTFA 1800
 TTTTACAGAG CTAGCCCAAG GCACAAGATG AAAATTATTA AGTCGCTACA GAAGAACGGT 1860
 TCAGTTGTAG CCAATGACAG AGATGGAGTA AATGATGCAG TTGCTCTGAA GGCTGCAGAC 1920

5 ATTGAGGTTG CGATGGGCCA GACTGGTACA GATGTTTGCA AAGAGGCAGC AGACATGATC 1980
 CTAGTGGATG ATGATTTTCA AACCATATG TCTGCAATCG AAGAGGGTAA AGGGATTTAT 2040
 AATAACATTA AAAATTTTCGT TAGATTCCAG CTGAGCACGA GTATAGCAGC ATTAACCTTA 2100
 ATCTCATTTGG CTACATTAAT GAACCTTCCT AATCCTCTCA ATGCCATGCA GATTTTGTGG 2160
 ATCAATATTA TTATGGATGG ACCCCCAGCT CAGAGCCTTG GAGTAGAACC AGTGGATAAA 2220
 GATGTCATTC GTAAACCTCC TCGCAACTGG AAAGACAGCA TTTTGACTAA AAACCTTGATA 2280
 CTTAAATATC TTGTTTCATC AATAATCATT GTTTGTGGGA CTTTGTTTGT CTCTCTGGCGT 2340
 10 GAGCTACGAG ACAATGTGAT TACACCTCGA GACACAACAA TGACCTTCAC ATGCTTTGTG 2400
 TTTTGTGACA TGTTCAATGC ACTAAGTTCC AGATCCCAGA CCAAGTCTGT GTTTGAGATT 2460
 GGACTCTGCA GTAATAGAAT GTTTTGTCTAT GCAGTTCTTG GATCCATCAT GGGACAATTA 2520
 CTAGTTATTT ACTTTCCTCC GCTTCAGAAG GTTTTTCAGA CTGAGAGCCT AAGCATACTG 2580
 GATCTGTTGT TTCTTTTGGG TCTCACCTCA TCAGTGTGCA TAGTGGCAGA AATTATAAAG 2640
 AAGGTTGAAA GGACGAGGGA AAAGATCCAG AAGCATGTTA GTTCGACATC ATCATCTTTT 2700
 15 CTTGAAGTAT GA

SEQ ID NO:208 PAJ5 Protein sequence:

Protein Accession #: AAF27813

20 1 11 21 31 41 51
 MIPVLTSKKA SELPVSEVAS ILQADLQNL NKCEVSHRRA FHGWNEFDIS EDEPLWKYI 60
 SQFKNPLIML LLASAVISVL MHQFDDAVSI TVAILIVTV AFVQEURSEK SLEELSKLVP 120
 PECHCVREGK LEHTLARDLV PGDTVCLSVG DRVPADLRFL EAVDLSDIES SLTGETTPCS 180
 25 KVTAPQPAAT NGDLASRSNI AFMGTLVRCG KAKGVVIGTG ENSEFGEVFK MMQAEAPKT 240
 PLQKSMDDLK QQLSFYSFGI IGIIMLVGWL LGKDILEMFT ISVSLAVAAI PEGLPVVTV 300
 TLALGVMRMV KKRAIVKKLP IVETLGCCNV ICSDKTGTLT KNEMTVTHIF TSDGLHAEVT 360
 GVGVNQFGEV IVDGDVVHGF YNPAVSRIE AGCVCNDAVI RNNFLMGKPT EGALIALAMK 420
 MGLDLGLQQDY IRKAEYPPSS EQKWMMAVKCV HRTQQDRPEI CFMKGAYEQV IKYCTTYQSK 480
 30 GQTLTLTQQQ RDVYQOEKAR MGSAGLRVLA LASGPGLGQL TFLGLVLIID PPRTGVKEAV 540
 TFLIAGSVSI KMITGDSQET AVAIASRLGL YSKTSQSVSG EEIDAMDVQQ LSQIVPKVAV 600
 FYRASPRHKM KIIKSLQKNG SVVAMTGDGV NDAVALKAAD IGVAMGQTGT DVCKEAADMI 660
 LVDDDFQPTM SAIEEGKGYI NNINKFVRFO LSTSLAALT ILSLATLMNFP NPLNAMQILW 720
 INIIMDGPPA QSLGVEPVDK DVIRKPPRNW KDSILTKNLI LKILVSSIII VCGTLFVFWR 780
 35 ELRDNVITPR DTTMTFTCFV FDFMFNALSS RSQTKSVFEI GLCSNRMFCY AVLGSIMGQL 840
 LVYTFPPLQK VFQTESLSIL DLLFLLGLTS SVCIVAEIK KVERSREKIQ KHVSTSSSS 900
 LEV

SEQ ID NO:209 PAV4 VARIANT 1 DNA SEQUENCE

Nucleic Acid Accession #: N62096

Coding sequence: 1-1284 (underlined sequences correspond to start and stop codons)

45 1 11 21 31 41 51
 ATGGGCTACC AGAGGCAGGA GCCTGTCATC CCGCCGCAGA GAGGATTGCC TTATTCATATG 60
 AAGCAAGCTG GGTTCCTTTT GGGAAATATG CTTTATTCTT GGGTTTCATA TGTTACAGAC 120
 TTTTCCCTTG TTTTATTGAT AAAAGGAGGG GCCCTCTCTG GAACAGATAC CTACCACTCT 180
 TTGGTCAATA AAACCTTCGG CTTTCCAGGG TATCTGCTCC TCTCTGTCTT TCAGTTTGTG 240
 50 TATCCTTTTA TAGCAATGAT AAGTTACAAT ATAATAGCTG GAGATACTTT GAGCAAGATT 300
 TTTCAAAGAA TCCCAGGAGT TGATCTGAA AACGTGTTTA TTGGTCGCCA CTTCAATTAT 360
 GGACTTTCCA CAGTTACCTT TACTCTGCCT TTATCCTTGT ACCGAAATAT AGCAAAGCTT 420
 GGAAAGGTCT CCCTCATCTC TACAGGTTTA ACAACTCTGA TTCTTGGAAT TGTAAATGGCA 480
 AGGGCAATT CTCTGGGTCC ACACATACCA AAAACAGAAG ACGCTTGGGT ATTTGCAAG 540
 55 CCCAATGCCA TCAAGCGGT CGGGGTATG TCTTTTGCAT TTATTTGCCA CCATAACTCC 600
 TCTTTAGTTT ACAGTTCTCT AGAAGAACC ACAGTAGCTA AGTGGTCCCG CCTATCCAT 660
 ATGTCCATCG TGATTTCTGT ATTTATCTGT ATATTCTTTG CTACATGTGG ATACTTGACA 720
 TTTACTGGCT TCACCCAAGG GGACTTATTT GAAAATTACT GCAGAAATGA TGACCTGGTA 780
 ACATTGGGAA GATTTTGTTA TGGTGTCACT GTCATTTTGA CATACCTAT GGAATGCTTT 840
 60 GTGACAAAG AGGTAAATGC CAATGTGTTT TTTGGTGGGA ATCTTTCATC GGTTTCCAC 900
 ATTGTTGTAA CAGTGATGGT CATCACTGTA GCCACGCTTG TGTCATTGCT GATTGATTGC 960
 CTCGGGATAG TTCTAGAACT CAATGGTGTG CTCTGTGCAA CTCCCTCAT TTTTATCATT 1020
 CCATCAGCCT GTTATCTGAA ACTGTCTGAA GAACCAAGGA CACACTCCGA TAAGATTATG 1080
 TCTTGTGTCA TGCTTCCCAT TGGTGTGTG GTGATGGTTT TTGGATTCTG CATGGCTATT 1140
 65 ACAAATACTC AAGACTGCAC CCATGGGCAG GAAATGTCT ACTGCTTTCC TGACAATTTT 1200
 TCTCTCAAA ATACCTCAGA GTCTCATGTT CAGCAGACAA CACAACCTTC TACTTTAAAT 1260
 ATTAGTATCT TTCAACTCGA GTAA

SEQ ID NO:210 PAV4 Variant 1 Protein sequence:

Protein Accession #: none found

75 1 11 21 31 41 51
 MGYQRQEPVI PFQRLPYSM KQAGFPLGIL LLFWVSIVTD FSLVLLIKGG ALSGTDYQS 60
 LVNKTGFPFG YLLLSVLQFL YFFIAMISYN IIAGDTLSKV FORIPGVDPE NVFIGRHFII 120
 GLSTVTFTLP LSLYRNIAKL GKVSLISTGL TTLILGIVMA RAISLGHIP KTEDAWVFAK 180
 PNAIQAVGVM SFAFICHHNS FLVYSSLEEP TVAKWSRLIH MSIVISVFIC IFFATCGYLT 240
 FTGFTQGLDF ENYCRNDLLD TFGRCYGVV VILTYPMECF VTREVIANVF FGGNLSVVFH 300
 80 IVVTVMVITV ATLVSLLLDC LGIVLELNGV LCATPLIFII PSACYLKLSE EPRTHSKDIM 360
 SCVMLPIGAV VMVFGFVMAI TNTQDCTHGQ EMFYCFPDNF SLTNTSESHV QQTQLSTLN 420

ISIFQLE

SEQ ID NO:211 PAV4 VARIANT 2 DNA SEQUENCE

5 Nucleic Acid Accession #: N62096
Coding sequence: 1-1203 (underlined sequences correspond to start and stop codons)

```

10 1      11      21      31      41      51
    |      |      |      |      |      |
    ATGGGCTACC AGAGGCAGGA GCCTGTCATC CCGCCGCAGT TTTCCTTGT TTTATTGATA 60
    AAAAGGAGGG CCTCTCTGG AACAGATACC TACCAGTCTT TGGTCAATAA AACTTTCGGC 120
    TTCCAGGGT ATCTGCTCCT CTCTGTTCTT CAGTMTTGT ATCCTTTTAT AGCAATGATA 180
    AGTTACAATA TAATAGCTGG AGATACTTTG AGCAAAGTTT TTCAAAGAAT CCCAGGAGTT 240
    15 GATCCTGAAA ACGTGTATTAT TGGTCGCCAC TTCAATTATTG GACTTTCAC AGTTACCTTT 300
    ACTCTGCCTT TATCCTTGTA CCGAAATATA GCAAAGCTTG GAAAGGTCTC CCTCATCTCT 360
    ACAGGTTTAA CAACTCTGAT TCTTGGAAAT GTAATGGCAA GGGCAATTTC ACTGGGTCCA 420
    CACATACCAA AAACAGAAGA CGCTTGGGTA TTGCAAAGC CCAATGCCAT TCAAGCGGTC 480
    GGGGTATGT CTTTGTCAT TATTTGCCAC CATAACTCCT TCTTAGTTTA CAGTTCTCTA 540
    20 GAAGAACCCA CAGTAGCTAA GTGGTCCCGC CTATCCATA TGTCATCGT GATTTCGTGA 600
    TTTATCTGTA TATCTTTGC TACATGTGGA TACTTGACAT TTACTGGCTT CACCCAAGGG 660
    GACTTATTTG AAAATTAAGT CAGAAATGAT GACCTGGTAA CATTTTGGAAG ATTTTGTAT 720
    GGTGTCAGTG TCATTTTGAC ATACCCATATG GAATGCTTTG TGACAAGAGA GGTAAATGCC 780
    AATGTGTTTT TTGGTGGGAA TCTTTTCATCG GTTTTCCACA TTGTTGTAAAC AGTGATGGTC 840
    25 ATCACTGTAG CCACGCTTGT GTCAATGCTG ATTTGATTGCC TCGGGATAGT TCTAGAACTC 900
    AATGGTGTGC TCTGTGCAAC TCCCTTCAT TTTATCATTC CATCAGCCTG TTAGTCTGAAA 960
    CTGCTTGAAG AACCAAGGAC ACACCTCCGAT AAGATTATGT CTGTGTGCAT GCTTCCCATT 1020
    GGTGCTGTGG TGATGTTTTT TGGATTTCGT ATGGCTATTA CAAATACTCA AGACTGCACC 1080
    CATGGGCAGG AAATGTTCTA CTGCTTCTCT GACAAATTCT CTCTCACAAA TACCTCAGAG 1140
    30 TCTCATGTTT AGCAGACAAC ACAACTTTCT ACTTTAAATA TTAGTATCTT TCAACTCGAG 1200
    TAA
  
```

SEQ ID NO:212 PAV4 Variant 2 Protein sequence:

35 Protein Accession #: none found

```

40 1      11      21      31      41      51
    |      |      |      |      |      |
    MGYQRQEPVI PPQFSVLVLI KGGALSGTDT YQSLVNKTFG FPGYLLLSVL QFLYPIFIAMI 60
    SYNIIAGDTL SKVFQRIPIV DPENVFIRH FIIGLSTVTF TLPLSLYRNI AKLGKVSILIS 120
    TGLTTLILGI VMARATSLGP HLPKTEDAWV FAKPNAIQAV GVMSFAFICH HNSFLVYSSL 180
    EEPTVAKWSR LIHMSIVISV FICIFFATCG YLTFTGFTQG DLFFENYCRND DLVTFGRFCY 240
    45 GVTVILTYPM ECFVTVREVIA NVFFGGLNLS VFHIVVTVMV ITVATLVSLI IDCLGIVLEL 300
    NGVLCATPLI FIIPSACYLK LSEEPRTSD KIMSCVMLPI GAVVMVFGFV MAITNTQDCT 360
    HGQEMFYCFP DNFSLTNTSE SHVQQTQLS TLNISIFQLE
  
```

SEQ ID NO:213 PAV4 VARIANT 3 DNA SEQUENCE

50 Nucleic Acid Accession #: N62096
Coding sequence: 1-1140 (underlined sequences correspond to start and stop codons)

```

55 1      11      21      31      41      51
    |      |      |      |      |      |
    ATGGGCTACC AGAGGCAGGA GCCTGTCATC CCGCCGCAGG TCAATAAAAC TTTCGGCTTT 60
    CCAGGGTATC TGCTCTCTCT TGTCTTTCAG TTTTGTATC CTTTTATAGC AATGATAAGT 120
    TACAATATAA TAGCTGGAGA TACTTTGAGC AAAGTTTTC AAAGATCCC AGGAGTTGAT 180
    CCTGAAAACG TGTTTATTGG TCGCCACTTC ATTATTGGAC TTTCACAGT TACCTTTACT 240
    CTGCCTTTAT CCTGTGACCG AAATATAGCA AAGCTTGGAA AGGTCTCCCT CATCTCTACA 300
    60 GGTTTAACAA CTCTGATTCT TGGAAATTGTA ATGGCAAGGG CAATTTCACT GGGTCCACAC 360
    ATACCAAAAA CAGAAGACGC TTGGGTATTG GCAAAGCCCA ATGCCATTCA AGCGGTCCGG 420
    GTTATGTCTT TTGCAATTAT TTGCCACCAT AACTCCTTCT TAGTTTACAG TTCTCTAGAA 480
    GAACCCACAG TAGCTAAGTG GTCCCGCCTT ATCCATATGT CCATCGTGAT TTCTGTATTT 540
    ATCTGTATAT TCTTTGCTAC ATGTGGATAC TTGACATTTA CTGGCTTCAC CCAAGGGGAC 600
    65 TTATTTGAAA ATTACTGCAG AAATGATGAC CTGGTAACAT TTGGAAGATT TTGTTATGGT 660
    GTCACTGTCA TTTTGACATA CCCTATGGAA TGCTTTGTGA CAAGAGAGGT AATTGCCAAT 720
    GTGTTTTTTG GTGGGAATCT TTCACTCGGT TTCCACATTG TTGTAACAGT GATGGTCATC 780
    ACTGTAGCCA CGCTTGCTC ATTGCTGATT GATTGCCTCG GGATAGTTCT AGAACTCAAT 840
    70 GGTGTGCTCT GTGCACTCC CCTCAATTTT ATCATTTCCAT CAGCCTGTTA TCTGAAACTG 900
    TCTGAAGAAC CAAGGACACA CTCCGATAAG ATTATGTCTT GTGTCATGCT TCCCATTGGT 960
    GCTGTGGTGA TGGTTTTTGG ATTCGTCATG GCTATTACAA ATACTCAAGA CTGCACCCAT 1020
    GGGCAGGAAA TGTTCTACTG CTTTCTCTGAC AATTTCTCTC TCACAATAC CTCAGAGTCT 1080
    CATGTTTCAGC AGACAACACA ACTTTCTACT TTAAATATTA GTATCTTTCA ACTCGAGTAA
  
```

SEQ ID NO:214 PAV4 Variant 3 Protein sequence:

75 Protein Accession #: none found

```

80 1      11      21      31      41      51
    |      |      |      |      |      |
  
```

MGYQRQEPVI PPQVNTKFGF PGYLLLSVLQ FLYPFIAMIS YNIIAGDTLS KVFQRI PGVD 60
 PENVFIGRHF IIGLSVTFIT LPLSLYRNIA KLKVSLSLIST GLTTLILGIV MARAISLGPH 120
 IPKTEDAWVF AKENAIQAVG VMSFAFICHH NSFLVYSSLE EPTVAKWSRL IHMSIVISVF 180
 ICIFFATCGY LTFTGTFQGD LFENYCRNDD LVTFRFCYCY VTIVILTYPME CFVTREVIAN 240
 VFEGGNLSSV FHIVVTVMVI TVATLVSLLI DCLGIVLELN GVLCAITPLIF IIPSACYLKL 300
 SEEPRTTHSDK IMSCVMLPIG AVVMVFGFVM AITNTQDCTH GQEMFYCFPD NFSLTNTSES 360
 HVQQTQLST LNISIFQLE

10

SEQ ID NO:215 PAV4 VARIANT 4 DNA SEQUENCE:

Nucleic Acid Accession #: N62096
 Coding sequence: 1-1389 (underlined sequences correspond to start and stop codons)

15

1 11 21 31 41 51
 ATGGGCTACC AGAGGCAGGA GCCTGTCATC CCGCCGCAGA GAGATTTAGA TGACAGAGAA 60
 ACCCTTGTTT CTGAACATGA GTATAAAGAG AAAACCTGTC AGTCTGCTGC TCTTTTAAAT 120
 GTTGTCAACT CGATTATAGG ATCTGGTATA ATAGGATTGC CTTATTCAAT GAAGCAAGCT 180
 GGGTTTCCCT TGGGAATATT GCTTTTATTC TGGGTTTCAT ATGTTACAGA CTTTTCCTTT 240
 GTTTTATTGA TAAAGAGGAG GGCCCTCTCT GGAACAGATA CCTACCAGTC TTTGGTCAAT 300
 AAAACTTTCG GCTTTCCAGG GTATCTGCTC CTCTCTGTTT TTCAGTTTTC GTATCCTTTT 360
 ATAGCAATGA TAAGTTACAA TATAATAGCT GGAGATACTT TGAGCAAAGT TTTTCAAAGA 420
 ATCCCAAGGAG TTGATCCTGA AAACGTGTTT ATTGGTCGCC ACTTCATTAT TGGACTTTCC 480
 ACAGTTACCT TTACTCTGCC TTTATCCTTG TACCGAAATA TAGCAAAGCT TGGAAAGGTC 540
 TCCCTCATCT CTACAGGTTT AACAACTCTG ATCTCTGGAA TTGTAATGGC AAGGGCAATT 600
 TCATCGGCTC CACACATACC AAAAACAGAA GACGCTGGG TATTGTGCAA GCCCAATGCC 660
 ATTCAAGCGG TCGGGGTTAT GTCTTTTGCA TTTATTTGCC ACCATAACTC CTCTTAGTTC 720
 TACAGTTCTC TAGAAGAACC CACAGTAGCT AAGTGGTCCC GCCTTATCCA TATGTCCATC 780
 GTGATTTCTG TATTATCTGT TATATCTTTC GCTACATGTG GATACCTGAC ATTACTTGGC 840
 TTCACCCAAAG GGGACTTATT TGAAAATTAC TGCAGAAATG ATGACCTGGT AACATTGGGA 900
 AGATTTTGTT ATGGTGTAC TGTCATTTTG ACATACCCTA TGGAAATGCT TGTGACAAGA 960
 GAGGTAATTG CCAATGTGTT TTTTGGTGGG AATCTTTCAT CGGTTTTCCT CATTGTGTGA 1020
 ACAGTGATGG TCATCAGTGT AGCCACGCTT GTGTCTATGC TGATTGATTG CCTCGGGATA 1080
 GTTCTAGAAC TCAATGGTGT GCTCTGTGCA ACTCCCTCA TTTTATCAT TCCATCAGCC 1140
 TGTTCATCTGA AACTGTCTGA AGAACCAAGG ACACACTCCG ATAAGATTAT GTCTTGTGTC 1200
 ATGCTTCCCA TTGGTGCTGT GGTGATGGTT TTTGGATTTC TCATGGCTAT TACAAATACT 1260
 CAAGACTGGA CCCATGGGCA GGAATGTTC TACTGCTTTC CTGACAATT TCTCTCACA 1320
 AATACCTCAG AGTCTCATGT TCAGCAGACA ACACAACCTT TACTTTTAAA TATTAGTATC 1380
 TTTCAATGA

40

SEQ ID NO:216 PAV4 Variant 4 Protein sequence:

Protein Accession #: none found

45

1 11 21 31 41 51
 MGYQRQEPVI PPQRDLDDRE TLVSEHEYKE KTCQSAALFN VVNSIIGSGI IGLPYSMKQA 60
 GFPLGILLF WVSYYTDFSL VLLIKGGALS GTDTYQSLVN KTFGFPYGL LSVLQFLYPP 120
 IAMISYNIIA GDTLSKVFQR IPGVDPENVF IGRHFIIGLS TVTFTLPLSL YRNIAKLGKV 180
 SLISTGLTTL ILGIVMARAI SLGPHIPKTE DAWVFAKPNA IQAVGVMSFA FICHNSFLV 240
 YSSLEBPTVA KWSRLHMTF VISVFCIFF ATCGYLFTFG FTQGDLEFNY RCHNDLVTFG 300
 RFCYGVTVIL TYPMCEFVTR EVIANVFFGG NLSSVFHIVV TVMVITVATL VSLIDCLGI 360
 VLELNGVLC TPLIFIPSA CYLKLSEEP THSKIMSCV MLPIGAVVMV FGFVMAITNT 420
 QDCTHGEQEM YCFPDNFSLT NTSESHVQQT TQLSTLNI SI FQ

55

SEQ ID NO:217 PAV9 DNA SEQUENCE

Nucleic Acid Accession #: NM_017636
 Coding sequence: 1-3501 (underlined sequences correspond to start and stop codons)

60

1 11 21 31 41 51
 ATGGAGGATG CCTTCGGGGC AGCCGTGGTG ACCGTGTGGG ACAGCGATGC ACACACCACG 60
 GAGAGGCCA CCGATGCCTA CGGAGAGCTG GACTTCACGG GGGCCGGCCG CAAGCACAGC 120
 AATTTCCTCC GGCTCTCTGA CCGAACGGAT CCAGCTGCAG TTTATAGTCT GGTCACACGC 180
 ACATGGGGCT TCCGTGCCCC GAACCTGGTG GTGTCACTGC TGGGGGGATC GGGGGGCCCC 240
 GTCTCTCAGA CTTGGCTGCA GGACCTGCTG CGTCTGGGGC TGGTGCGGGC TGCCACAGGC 300
 ACAGGAGCCT GGATGTCTAC TGGGGGTCTG CACACGGGCA TCGGCCGGCA TGTGGGTGTG 360
 GCTGTACGGG ACCATCAGAT GGCCAGCACT GGGGGCACCA AGGTGGTGGG CATGGGTGTG 420
 GCCCCTGGG GTGTGGTCCG GAATAGAGAC ACCCTCATCA ACCCAAGGGC CTCGTTCCCT 480
 GCGAGGTACC GGTGGCGCGG TGACCCGGAG GACGGGGTCC AGTTTCCCTT GGACTACAAC 540
 TACTCGGCCT TCTTCTGGT GGACGACGGC ACACACGGCT GCCTGGGGGG CGAGAACC GC 600
 TTCCGCTTGC GCCTGGAGTC CTACATCTCA CAGCAGAAGA CGGGCGTGGG AGGGACTGGA 660
 ATTGACATCC CTGTCTGTCT CCTCTGATT GATGGTGTAT AGAAGATGTT GACGCGAATA 720
 GAGAACGCCA CCCAGGCTCA GCTCCCATGT CTCTCTGGT CTGGCTCAGG GGGAGCTGCG 780
 GACTGCCTGG CGGAGACCTT GGAAGACACT CTGGCCCCAG GGAGTGGGGG AGCCAGGCAA 840
 GGCGAAGCCC GAGATCGAAT CAGGCGTTTC TTTCCCAAAG GGGACCTTGA GGTCTGCAG 900
 GLCCAGGTGG AGAGGATTAT GACCCGGAAG GAGCTCTTGA CAGTCTATTC TTCTGAGGAT 960

80

	GGGTCTGAGG	AATTCGAGAC	CATAGTTTTC	AAGGCCCTTG	TGAAGGCCTG	TGGGAGCTCG	1020
	GAGGCCTCAG	CCTACCTTGA	TGAGCTGCGT	TTGGCTGTGG	CTTGAACCCG	CGTGGACAT	1080
	GCCCAGAGTG	AACCTCTTTC	GGGGGACATC	CAATGGCGGT	CCTTCCATCT	CGAAGCTTCC	1140
5	CTCATGGACG	CCCTGCTGAA	TGACCGGCCT	GAGTTCGTGC	GCTTGCTCAT	TTCCACCGGC	1200
	CTCAGCCTGG	GCCACTTCTT	GACCCCGATG	CGCCTGGCCC	AACTCTACAG	CGCGCGCCCC	1260
	TCCAACTCGC	TCAATCCCAA	CCTTTTGGAC	CAGGCGTCCC	ACAGCGCAGG	CACCAAAGCC	1320
	CCAGCCCTAA	AAGGGGGAGC	TGCGGAGCTC	CGGCCCCCTG	ACGTGGGGCA	TGTGCTGAGG	1380
	ATGCTGTGTT	GGAGAGTGTG	CGCGCCGAGG	TACCCCTCCG	GGGGCGCCTG	GGACCCCTAC	1440
10	CCAGGCCAGG	GCTTCGGGGA	GAGCATGTAT	CTGCTCTCGG	ACAAGGCCAC	CTCGCCGCTC	1500
	TCGCTGGATG	CTGGCCTCGG	GCAGGCCCCC	TGGAGCGACC	TGCTTCTTTG	GGCACTGTTC	1560
	CTGAACAGGG	CACAGATGGC	CATGTACTTC	TGGGAGATGG	GTTCCAATGC	AGTTTCTCTA	1620
	GCTCTTGGGG	CCTGTTTTCG	GCTCCGGGTG	ATGGCACGCC	TGGAGCCTGA	CGCTGAGGAG	1680
	GCAGCACGGA	GGAAAGACCT	GGCGTTCAAG	TTTGAGGGGA	TGGGCGTTGA	CCTCTTTGGC	1740
15	GAGTGCTATC	GCAGCAGTGA	GGTGAGGGCT	GCCCGCCTCC	TCCTCCGTGC	CTGCCCCGTC	1800
	TGGGGGGATG	CCACTTGCCCT	CCAGCTGGCC	ATGCAAGCTG	ACGCCCGTGC	CTTCTTTGCC	1860
	CAGGATGGGG	TACAGTCTCT	GCTGACACAG	AAGTGGTGGG	GAGATATGGC	CAGCACTACA	1920
	CCCATCTGGG	CCCTGGTTCCT	CGCCTTCTTT	TGCCCTCCAC	TCATCTACAC	CCGCCCTCATC	1980
	ACCTTCAGGA	AATCAGAAAG	GGAGCCACAC	CGGGAGGAGC	TAGAGTTTGA	CATGGATAGT	2040
20	GTCATTAATG	GGGAAGGGCC	TGTCGGGACG	SCGGACCCAG	CCGAGAAGAC	GCCGCTGGGG	2100
	GTCCCGCGCC	AGTCGGGTTC	TCCGGGTTCG	TGCGGGGGCC	GCTGCGGGGG	GCGCCGGTGC	2160
	CTACGCCGCT	GGTTCACACT	CTGGGGCGCG	CCGGTGACCA	TCTTCATGGG	CAACGTGGTC	2220
	AGCTACCTGC	TGTTCCTGCT	GCTTTTCTCG	CGGGTGCTGC	TCGTGGATT	CCAGCCGGCG	2280
	CCGCCCCGCT	CCCTGGAGCT	GCTGCTCTAT	TTCTGGGCTT	TCACGCTGCT	GTGCGAGGAA	2340
25	CTGCGCCAGG	GCCTGAGCGG	AGGCGGGGGC	AGCCTCGCCA	CGGGGGGGCC	CGGGCCTGGC	2400
	CATGCCCTAC	TGAGCCAGCG	CCTGCGCCTC	TACCTCGCCG	ACAGCTGGAA	CCAGTGCGAC	2460
	CTAGTGCCCT	TCACCTGCTT	CCCTCTGGGC	GTGGGCTGCC	GGCTGACCCC	GGGTTTGTAC	2520
	CACCTGGGCC	GCATGTTCCT	CTGCATCGAC	TTTATGGTTT	TCACGGTGCG	GCTGCTTCAC	2580
	ATCTTCACGG	TCAACAAACA	GCTGGGGCCC	AAGATCGTCA	TCGTGAGCAA	GATGATGAAG	2640
30	GACGTGTCTT	TCTTCCTCTT	CTTCCTCGCG	GTGTGGCTGG	TAGCCCTATG	CGTGGCCACG	2700
	GAGGGGCTCC	TGAGGCCACG	GGACAGTGAC	TTCCCAAGTA	TCCTGCGCCG	CGTCTTCTAC	2760
	CGTCCCTACC	TGCAGATCTT	CGGGCAGATT	CCCAGGAGG	ACATGACAGT	GGCCCTCATG	2820
	GAGCACAGCA	ACTGCTCGTC	GGAGCCCGGC	TTCTGGGCAC	ACCCTCCTGG	GGCCAGGCGG	2880
35	GGCACCTGCG	TCCTCCAGTA	TGCCAACTGG	CTGGTGGTGC	TGCTCCTCGT	CATCTTCCTG	2940
	CTCTGTGGCA	ACATCTCTGT	GGTCAACTTG	CTCATTGCCA	TGTTCAAGTA	CACATTCGGC	3000
	AAAGTACAGG	GCAACAGCGA	TCTCTACTGG	AAGGCGCAGC	GTTACCGCCT	CATCCGGGAA	3060
	TTCCACTCTC	GGCCCGCGCT	GGCCCGCCCC	TTTATCGTCA	TCTCCCACTT	GCGCCTCCTG	3120
	CTCAGGCAAT	TGTGACGGCG	ACCCCGGAGC	CCCCAGCCGT	CCTCCCGGCG	CCTCGAGCAT	3180
40	TTCCGGGTTT	ACCTCTCTAA	GGAAAGCTGC	TAACGTGGGA	ATCGGTGCAT	ATCGGTGCAT	3240
	AAGGAGAACT	TTCTGCTGGC	ACGCGCTAGG	GACAAGCGGG	AGAGCGACTC	CGAGCGTCTG	3300
	AAGCGCACGT	CCCAAGAGGT	GGACTTGGCA	CTGAAACAGC	TGGGACACAT	CCGCGAGTAC	3360
	GAACAGCGCC	TGAAAGTGCT	GGAGCGGGAG	GTCCAGCAGT	GTAGCCGCGT	CTGGGGTGGG	3420
	GTGGCCGAGG	CCCTGAGCCG	CTCTGCCTTG	CTGCCCCCAG	GTGGGCGGCC	ACCCCTGTAC	3480
	CTGCCTGGGT	CCAAAGACTG	A				

SEQ ID NO:218 PAV9 Protein sequence:

Protein Accession #: none found

	1	11	21	31	41	51	
50	MEDAFGAADV	TVWDSDAHTT	EKPTDAYGEL	DFTGAGRKHS	NFLRLSDRTD	PAAVYSLVTR	60
	TWGFRAPNLV	VSVLGGSGGP	VLQIWLQDLL	RRGLVRAAQS	TGAMIVTGGL	HTGIGRHVGV	120
	AVRDHQMST	GGTKVVMAMV	APWGVVRNRD	TLINPKGSFP	ARYRWGRDPE	DGVQFPLDYN	180
55	YSAFFLVDDG	THGCLGAGTA	FRLRLSEYIS	QKQTVGGTGG	IDIPVLLLLI	DGDEKMLTRI	240
	ENATQAQLPC	LLVAGSGGAA	DCLAETLEDT	LAPGSGGARQ	GEARDRIIRF	FPKGDLEVLQ	300
	AQVERIMTRK	ELLTVVSSSD	GSEEFETIVL	KALVKACGSS	EASAYLDELRL	LAVAWNVRDI	360
	AQSELFGRDI	QWRSFHLLEAS	IMDALLNDRP	EFVRLILSHG	LSLGHFLTPM	RLAQLYSAAP	420
	SNSLIRNLLD	QASHSAGTKA	PALKGGAAEL	RPPDVGHVLR	MLLGKMCAPR	YPSGGAWDPH	480
60	PGQFGESMY	LLSDKATSP	SLDAGLQAP	WSDLLWALL	LNRAQMAMYF	WEMGSNAVSS	540
	ALGACLLLRV	MARLEPDAEE	AARRKDLAFK	FEGMGVDFLG	ECYRSSEVRA	ARLLLRRCPL	600
	WGDATCLQLA	MQADARAFFA	QDGVQSLLTQ	KWWDMASTT	PIWALVLAFF	CPPLIYTRLI	660
	TRFKSEEEPT	REELEPDMDS	VINGEGPVTG	ADPAEKTPLG	VPRQSGRPGC	CGGRCGGRR	720
	LRWFHFWGA	PVTIFMGNV	SYLLFLLFFS	RVLLVDFQPA	PPGSELELLY	FWAFTLLCEE	780
65	LRQLSGGGG	SLASGGPGPG	HASLSQRLRL	YLADSWNQCD	LVALTCFLLG	VGCRLLTPGLY	840
	HLGRTVLCID	FMVFTVRLHL	IFTVNKQLGP	KIVIVSKMMK	DVFFFLFFLG	VWLVAYGVAT	900
	EGLLRPRDSD	FPSILRRVFY	RPYLQIFGQI	PQEDMDVALM	EHSNCSSEPG	FWAHPPGAQA	960
	GTCVSQYANW	LVVLLLVIFL	LVANILLVNL	LIAMFSYTFG	KVQNSDLVW	KAQRYRLIRE	1020
	FHSRPAALAP	FIVISHLRLL	LRQLCRRFRS	PQSSSPALEH	FRVYLSKEAE	RKLLTWESVH	1080
70	KENFLARAR	DKRESDSERL	KRTSQKVDLA	LKQLGHIREY	EQRLKVLERE	VQQCSRVLGW	1140
	VAEALSRSAL	LPPGGPPPPD	LPGSKD				

SEQ ID NO:219 PBF1 DNA SEQUENCE

Nucleic Acid Accession #: AA054237

Coding sequence: 1-894 (underlined sequences correspond to start and stop codons)

	1	11	21	31	41	51	
75	ATGGAGCCCG	GGGCGCTCGT	CACGGCGCTC	AGCCTCGGCC	TCAGCCTGTG	CTCCCTGGGG	60
	CTGCTCGTCA	CGGCCATCTT	CACCGACCAC	TGGTACGAGA	CCGACCCCGG	GGGCCACAAG	120
80	GAGAGCTGCG	AGCGCAGCCG	CGCGGGCGCC	GACCCCGCGG	ACCAGAAGAA	CCGCCGTGAT	180

5
10
15
20
25

```

CCGCTGTGCG ACCTGCCGCT GCGGGACTCG CCCCCGCTGG GCGCGCGGCT GCTCCCGGGC 240
GGCCCGGGGC GCGCCGACCC CGAGTCCTGG CGCTCGCTCC TGGGGCTCGG CGGGCTGGAC 300
GCCGAGTGCG GCCGGCCCTT CTTCGCCACC TACTCGGGCC TCTGGAGGAA GTGCTACTTC 360
CTGGGCATCG ACCGGGACAT CGACACCCCT ATCCTGAAAG GTATTGCGCA GCGATGCACG 420
GCCATCAAGT ACCACTTTTC TCAGCCCATC CGCTTGCGAA ACATTCCCTT TAATTAAACC 480
AAGACCATAC AGCAAGATGA GTGGCACCTG CTTCATTTAA GAAGAATCAC TGCTGGCTTC 540
CTCGGCATGG CCGTAGCCGT CTTCTCTGCG GGCTGCATTC TGGCCACAGT CAGTTTCTTC 600
TGGGAGGAGA GCTTGACCCA GCACGTGGCT GGACTCCTGT TCCTCATGAC AGGGATATT 660
TGCAACATTT CCCTCTGTAC TTATGCCGCC AGTATCTCGT ATGATTGAA CCGGCTCCCA 720
AAGCTAATTT ATAGCCTGCC TGCTGATGTG GAACATGGTT ACAGCTGGTC CATCTTTTGC 780
GCCTGGTGCA GTTTAGCCTT TATTGTGGCA GCTGGAGTCT TCTGCATCGC TTATCCGTTT 840
ATTAGCCGGA CCAAGATTGC ACAGCTAAAG TCTGGCAGAG ACTCCACGGT ATGA

```

SEQ ID NO:220 PBF1 Protein sequence:

Protein Accession #: none found

20
25

```

1 11 21 31 41 51
| | | | |
MEPRALVTAL SLGLSLCSLG LLVTAIFTDH WYETDPRRHK ESCERSRAGA DPPDQKNRLM 60
PLSHLPLRDS PPLGRRLLPG GPGRADPEWS RSLGLGLGGL AECGRPLFAT YSGLWRKCYF 120
LGDDRDDTL ILKGIARQCT AIKYHFSQPI RLRNIFNLAT KTIQDEWHL LHLRRITAGF 180
LGMVAVLLC GCIVATVSFF WEESLTQHVA GLLFLMTGIF CTISLCTYAA SISYDLNRLP 240
KLITYSLPADV EHGYSWSIFC AWCSLGFIVA AGGLCIAYPF ISRTKIAQLK SGRDSTV

```

SEQ ID NO:221 PC14 DNA SEQUENCE

Nucleic Acid Accession #: NM_016570

Coding sequence: 1- 1134 (underlined sequences correspond to start and stop codons)

30
35
40
45
50

```

1 11 21 31 41 51
| | | | |
ATGAGGCGAC TGAATCGGAA AAAAACTTTA AGTTTGGTAA AAGAGTTGGA TGCCTTTCCG 60
AAGGTTCCCTG AGAGCTATGT AGAGACTTCA GCCAGTGGAG GTACAGTTTC TCTAATAGCA 120
TTTACAACCTA TGGCTTTTATT AACCATAATG GAATTCTCAG TATATCAAGA TACATGGATG 180
AAGTATGAAT ACCAAGTAGA CAAGGATTTT TCTAGCAAAT TAAGAATTAA TATAGATATT 240
ACTGTTGCCA TGAAGTGTCA ATATGTTGGA GCGGATGTAT TGGATTAGC AGAAACAATG 300
GTTGCATCTG CAGATGGTTT AGTTTATGAA CCAACAGTAT TTGATCTTTC ACCACAGCAG 360
AAAGAGTGGC AGAGGATGCT GCAGCTGATT CAGAGTAGGC TACAAGAAGA GCATTCACCT 420
CAAGATGTGA TATTTAAAG TGTCTTTAAA AGTACATCAA CAGCTCTTCC ACCAAGAGAA 480
GATGATTTCAT CACAGTCTCC AAATGCATGC AGAATTCATG GCCATCTATA TGTCATAAAA 540
GTAGCAGGGA ATTTTCACAT AACAGTGGGC AAGGCAATTC CACATCCTCG TGGTCATGCA 600
CAATTGGCAG CACTTGTCAA CCATGAATCT TACAATTTT CTCTAGAAAT AGATCATTTG 660
TCTTTTGGAG AGCTTGTTC AGCAATTATT AATCCTTTAG ATGGAAGTGA AAAAAATTGCT 720
ATAGATCACA ACCAGATGTT CCAATATTTT ATTACAGTTG TGCCAAACAA ACTACATACA 780
TATAAAATAT CAGCAGACAC CCATCAGTTT TCTGTGACAG AAAGGGAACG TATCATTAAC 840
CATGCTGCAG GCAGCCATGG AGTCTCTGGG ATATTTATGA AATATGATCT CAGTTCTCTT 900
ATGGTGACAG TTACTGAGGA GCACATGCCA TTCTGGCAGT TTTTGTGAAG ACTCTGTGGT 960
ATTGTTGGAG GAATCTTTTC AACCAAGGC ATGTTACATG GAATTGGAAA ATTTATAGTT 1020
GAAATAATTT GCTGTCTGTT CAGACTTGGG TCCTATAAAC CTGTCAATTC TGTTCCTTTT 1080
GAGGATGGCC ACACAGACAA CCACTTACCT CTTTATAGAA ATAATACACA TGA

```

SEQ ID NO:222 PC14 Protein sequence:

Protein Accession #: NP_057654

55
60
65

```

1 11 21 31 41 51
| | | | |
MRLNRKKTLL SLVKELDAFP KVPESVETS ASGGTVSLIA FTTMALLTIM EFSVYQDTWM 60
KYEYEVKDF SSKLRINIDI TVAMKQYVG ADVLDLAETM VASADGLVYE PTVFDLSPQQ 120
KEWQRMQLLI QSLRQEEHSL QDVIFKSAFK STSTALPPRE DDSSQSPNAC RIHGHLYVNK 180
VAGNFHITVG KAIHPHPRGHA HLAALVNHES YNFSHRIDHL SFGELVPAII NPLDGTAKIA 240
IDHNQMFQYF ITVVPRLKHT YKISADTHQF SVTERERIIN HAAGSHGVSG IFMKYDLSSL 300
MVTVTTEHMP FWQFFVRLCG IVGGIFSTTG MLHGIGKFIV EIIICFRFLG SYKPVNSVFP 360
EDGHTDNLHP LLENNTHT

```

SEQ ID NO:223 PEZ3 DNA SEQUENCE

Nucleic Acid Accession #: NM_001935.1

Coding sequence: 76-2301 (underlined sequences correspond to start and stop codons)

75
80

```

1 11 21 31 41 51
| | | | |
CGCGCGTCTC CGCGGCCCGC GTGACTTCTG CTTGCGCTCC TTCTCTGAAC GCTCACTTCC 60
GAGGAGACGC CGACGATGAA GACACCGTGG AAGATTCTTC TGGGACTGCT GGGTGTGTCT 120
GCCGTTGTCA CCATCATCAC CGTGCCCGTG GTTCTGCTGA ACAAAAGGCAC AGATGATGCT 180
ACAGCTGACA GTCCGAAAC TTACACTCTA ACTGATTACT TAAAAAATAC TTATAGACTG 240
AAGTTATACT CCTTAAAGATG GATTTCAGAT CATGAATATC TCTACAAACA AGAAAATAAT 300

```

5 ATCTTGGTAT TCAATGCTGA ATATGGAAAC AGCTCAGTTT TCTTGGAGAA CAGTACATTT 360
 GATGAGTTTG GACATCTCTAT CAATGATTAT TCAATATCTC CTGATGGGCA GTTTATTCTC 420
 TTAGAATACA ACTACGTGAA GCAATGGAGG CATTCCTACA CAGCTTCATA TGACATTTAT 480
 GATTTTAAATA AAAGGCAGCT GATTACAGAA GAGAGGATTC CAAACAACAC ACAGTGGGTC 540
 ACATGGTCAC CAGTGGGTCA TAAATTGGCA TATGTTTGGG ACAATGACAT TTATGTTAAA 600
 ATTGAACCAA ATTTACCAAG TTACAGAATC ACATGGACGG GGAAGAAGA TATAATATAT 660
 AATGGGAATA CTGACTGGGT TTATGAAGAG GAAGTCTTCA GTGCCTACTC TGCTCTGTGG 720
 TGGTCTCCAA ACGGCACTTT TTAGCATAT GCCCAATTTA ACGACACAGA AGTCCCACTT 780
 10 ATTGAATACT CCTTCTACTC TGATGAGTCA CTGCAGTACC CAAAGACTGT ACGGTTTCCA 840
 TATCCAAAGG CAGGAGCTGT GAATCCAAC TGAAGTTCT TTGTTGTAAA TACAGACTCT 900
 CTCAGCTCAG TCACCAATGC AACTTCCATA CAAATCAGT CTCCTGCTTC TATGTTGATA 960
 GGGGATCACT ACTTGTGTGA TGTGACATGG GCAACACAAG AAAGAATTTC TTTGCAGTGG 1020
 CTCAGGAGGA TTCAGAACTA TTCGGTCATG GATATTTGTG ACTATGATGA ATCCAGTGGG 1080
 15 AGATGGAAT GCTTAGTGGC ACGGCAACAC ATTGAATGA GTACTACTGG CTGGGTTGGA 1140
 AGATTTAGGC CTTCAGAAC CTATTTTACC CTGTATGGTA ATAGCTTCTA CAAGATCATC 1200
 AGCAATGAAG AAGGTTACAG ACACATTTGC TATTTCCAAA TAGATAAAAA AGACTGCACA 1260
 TTTATTACAA AAGGCACCTG GGAAGTCAATC GGGATAGAAG CTCTAACCCAG TGATATCTA 1320
 TACTACATTA GTAATGAATA TAAAGGAATG CCAGGAGGAA GGAATCTTTA TAAATCCAA 1380
 20 CTTATTGACT ATACAAAAGT GACATGCCTC AGTTGTGAGC TGAATCCGGA AAGGTGTCTCAG 1440
 TACTATTCTG TGCTATTCAG TAAAGAGGCG AAGTATTATC AGCTGAGATG TTCCGGTCTC 1500
 GGTCTGCCCC TCTATCTCTC ACACAGCAGC GTGAATGATA AAGGGCTGAG AGTCTGGGAA 1560
 GACAAATTCAG CTTTGGATTA AATGCTGCGA AATGTCAGG TGCCCTCCAA AAAACTGGAC 1620
 TTCATTATTT TGAATGAAAC AAAATTTTGG TATCAGATGA TCTTGCCCTC TCAATTTGAT 1680
 25 AAATCCAAAG AATATCCTCT ACTATTAGAT GTGTATGCGA GCCCATGTAG TCAAAAAGCA 1740
 GACACTGTCT TCAGACTGAA CTGGGCCACT TACCTTGCAA GCACAGAAAA CATTATAGTA 1800
 GCTAGCTTTG ATGGCAGAGG AAGTGGTTAC CAAGGAGATA AGATCATGCA TGCAATCAAC 1860
 AGAAGACTGG GAACATTTGA AGTTGAAGAT CAAATGAAG CAGCCAGACA ATTTTCAAAA 1920
 ATGGGATTTG TGGACAACAA ACGAATTGCA ATTTGGGGCT GGTACATATG AGGGTACGTA 1980
 30 ACCTCAATGG TCTGGGATG GGGAAAGTGG GTGTTCAAGT GTGGAATAGC CGTGGCGCCT 2040
 GTATCCCGGT GGGAGTACTA TGACTCAGTG TACACAGAAC GTTACATGGG TCTCCCAACT 2100
 CCAGAAGACA ACCTTGACCA TTACAGAAAT TCAACAGTCA TGAGCAGAGC TGAATATTTT 2160
 AAACAAGTTG AGTACCTCCT TATTCATGGA ACAGCAGATG ATAACGTTCA CTTTCAGCAG 2220
 TCAGCTCAGA TCTCCAAAGC CTTGGTTCAT GTTGGAGTGG ATTTCCAGGC AATGTGGTAT 2280
 35 ACTGATGAAG ACCATGGAAT AGCTAGCAGC ACAGCACACC AACATATATA TACCCACATG 2340
 AGCCACTTCA TAAACAATG TTTCTCTTTA CCTTAGCACC TCAAAATACC ATGCCATTTA 2400
 AAGCTTATTA AAATCATTTT TTGTTTTCAT TATCTCAAAA CTGCACTGTC AAGATGATGA 2460
 TGATCTTTAA AATACACACT CAAATCAAGA AACTTAAGGT TACCTTTGTT CCCAAATTTT 2520
 ATACCTATCA TCTTAAGTAG GGACTTCTGT CTTCACAA CA GATTATTACC TTACAGAAAGT 2580
 40 TGAATATATC CGGTCCGGTT TTATTGTTTA AAATCATTTT TGCATCAGCT GCTGAAACAA 2640
 CAAATAGGAA TTGTTTATTT GGAGGCTTTG CATAGATTCC CTGAGCAGGA TTTTAATCTT 2700
 TTTCTAACTG GACTGGTTCA AATGTTGTTT TCTTCTTTAA AGGGATGGCA AGATGTGGGC 2760
 AGTGAATGTA CTAGGGCAGG GACAGGATAA GAGGGATTAG GGAGAGAAGA TAGCAGGGCA 2820
 45 TGCTGGGAA CCGAGTCCA AGCATACCAA CACGAGCAGG CTACTGTCTG CTCCTCTCGG 2880
 AGAAGAGCTG TTCACCACGA GACTGGCACA GTTTTCTGAG AAAGACTATT CAAACAGTCT 2940
 CAGGAAATCA AATATCGAAA GCATGACTT CTAAATTAAC CACAGCAGTT GAAAGACTCC 3000
 AAAGAAATGT AAGGGAAACT GCCAGCAACG CAGCCCCAG GTGCCAGTTA TGGCTATAGG 3060
 TGCTACAAAA ACACAGCAAG GGTGATGGGA AAGCATTGTA AATGTGCTTT TAAAAAATA 3120
 TACTGATGTT CCTAGTGAAA GAGGCAGCTT GAAACTGAGA TGTGAACACA TCAGCTTGCC 3180
 50 CTGTTAAAG ATGAATAATAT TTGATACACA AATCTTAAC TGAAGGAGTC CTTGCATCAA 3240
 TTTTCTTAT TTCAATTTCT TGAGTGTCTT AATTAAGAATA ATATTTTAA TCTCTGGAC 3300
 TCATTTTAAA AAATGGAACA TAAATACAA TGTATGTAT TATTATCCC ATTCTACATA 3360
 CTATGGAATT TCTCCAGTC ATTTAATAAA TGTGCCTTCA TTTTTC

55 **SEQ ID NO:224 PEZ3 Protein sequence:**
 Protein Accession #: NP_001926.1

60 1 11 21 31 41 51
 MKTFWKILLG LLGAAALVTI ITVPVLLNKG GTDDATADSR KTYTLTDYLK NTYRLKLYSL 60
 RWISDHEYLY KOENNILVFN AEYGNSSVFL ENSTFDEFHG SINDYSISPD GQFILLEYN 120
 VKQWRHSYTA SYDIYDLNKR QLITEERIPN NTQVWTSVPV GHKLAYVWNN DIYVKIEPNL 180
 PSYRIWTGK EDIIYNGITD WYEEVEVFS YSALWWSPNG TFLAYAQFND TEVPLIEYSF 240
 65 YSDESLQYPK TVRVYPKAG AVNPTVKFFV VNTDSLSSVT NATSIQITAP ASMLIGDHYL 300
 CDVTWATQER ISLQWLRIQ NYSVMDICDY DESSGRWNCL VARQHIEMST TGWVGRFRES 360
 EPHFTLDGNS FYKILSNEEG YRHICYFQID KKDCTFITKG TWEVIGIEAL TSDYLYYISN 420
 EYKMPGGRN LYKIQILIDY KVTCLSCELN PERCQYYSVS FSKEAKYQYL RCGSPGLPLY 480
 TLHSSVNDKG LRVLEDNSAL DKMLQNVQMP SKKLDFIILN ETKFWYQML PPHFDKSKKY 540
 70 PLLLDVYAGP CSQKADTVFR LNWATYLAST ENIIVASF DG RSGSGVQDKI MHAINRRLGT 600
 FEVEDQIEAA RQFSKMGFVD NKRIAIWGS YGGYVTSMLV GSGSGVFKCG IAVAPVSRWE 660
 YYDSVYTER Y MGLPTPEDNL DHYRNSTVMS RAENFKQVEY LLIHGTADDN VHFQQAQIS 720
 KALVDVGVDF QAMWYTEDH GIASSTAHQH IYTHMSHFIEK QCFSLF

75 **SEQ ID NO:225 PBJ2 DNA SEQUENCE**
 Nucleic Acid Accession #: none found
 Coding sequence: 1-261 (underlined sequences correspond to start and stop codons)

80 1 11 21 31 41 51
 | | | | |

ATGGCTCTGG CGAAGGTGAG GGAGCCAAAC GCAAATGACA ATGCCATCAG AGTTGACAAC 60
 AGAAGTGTGA TTAAGTGCAG TGCTAACCCAG TGTTCCTGCG ATGAGGCAGA AAGTGAATCC 120
 AGAAACCCCTC AGGAGCTCTG GATGGGCTCG CTCCTCTTGA TGGGGGTCTT AGAAGCATGT 180
 GTGGAAATGA GGCTCTGTGC AGTCTGGTCC CTGAGAGATG ACAAGGAGCA GAGCCCCCAC 240
 CAGCCACAC TGGATGTCTA A

SEQ ID NO:226 PBJ2 Protein sequence:

Protein Accession #: none found

1 11 21 31 41 51
 | | | | |
 MALAKVREPN ANDNAIRVDN RSVIKVRANQ CSLHEAESES RNPQELWMGL LLLMGVLEAC 60
 VEMRPLSVWS LRDDKEQSPH QPTLDV

SEQ ID NO:227 PBM2 DNA SEQUENCE

Nucleic Acid Accession #: none found

Coding sequence: 1-462 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
ATGCCAAATG CTGAGTTAGA AGCAAAGAGC CTTGGAAGCA GTAAATGTTT AAAAACTGCT 60
 CTCATACCTTG CTGATATGTTG TGGATCAGCA AATATAGTCA GCCCTCTACT TGAGCAAAAT 120
 ATTGATGTAT CTCTCTCAAGA TCTGGACAGA CGGCCAGAGA GTATGCTGTT TCTAGTCATC 180
 ATCATGTGGA CCAGTTTGTG GGAAGACAAT CTTTCCATGG GCTGGGGGAA GCTAGAAGAT 240
 TTTATGGCTA TTGAAGAAGA AATGAAGAAG CACGGAAGTA CTCATGTGGG ATTCCAGAA 300
 AACCTGACTA ATGGTGCCGC TGCTGGCAAT GGTGATGATG GATTAAATCC TCCAAGGAAG 360
 AGCAGAACAC CTGAAAGCCA GCAATTTCCT GACACTGAGA ATGAAGAGTA TCACAGGTTT 420
 GTCAAGATC AGATAGTTGT AGATATGCGG CGTTATTTCT GA

SEQ ID NO:228 PBM2 Protein sequence:

Protein Accession #: none found

1 11 21 31 41 51
 | | | | |
 MPNAELEAKS LGSSKCLKTA LILAVCCGSA NIVSPLLEQN IDVSSQDLDR RPESMLFLVI 60
 IMWTSFVBDN LSMGWGKLED FMAIEEEMKK HGSTHVGFPE NLTNAGAAAGN GDDGLIPPRK 120
 SRTPEQQFP DTENEYHRF VKDQIVVDMR RYF

SEQ ID NO:229 PEZ2 DNA SEQUENCE

Nucleic Acid Accession #: NM_014253

Coding sequence: 65-8242 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 GACTGCTTGC ATTAAAGGAC TTCTCATCC TTTTTCAT GAAACTGAGC TTGCTTAATC 60
 AGAGATGGAG CAACTGACT GCAAACCTTA CCAGCCTCTA CCAAAAGTCA AGCATGAAAT 120
 GGATCTAGCT TACACGATT CTCTGATGA GAGTGAAGAT GGAAGAAAAC CAAGACAGTC 180
 ATACAACCTC AGGGAGACCC TGCACGAGTA TAACCAGGAG CTGAGGATGA ATTACAATAG 240
 CCAGAGTAGA AAGAGGAAAG AAGTAGAAAA ATCTACTCAA GAGATGGAAT TCTGTGAAAC 300
 CTCTCACACT CTGTGCTCTG GCTACCACAA AGACATGCAC AGCGTTTCTC GGCATGGCTA 360
 CCAGCTAGAG ATGGGATCTG ATGTGGACAC AGAGACAGAA GGTGCTGCCT CACCTGACCA 420
 TGCATAAGA ATGTGGATAA GGGGAATGAA ATCAGAGCAT AGTTCCTGTT TGTCCAGCCG 480
 GGCCAACCTC GCATTATCCT TGACTGACAC TGACCATGAA AGGAAGTCTG ATGGGGAAAA 540
 TGGTTTCAAA TTCTCTCCTG TTTGTTGTGA CATGGAGGCT CAAGCTGGGT CTAATCAAGA 600
 TGTGCAGAGC AGCCACACCA ACCAGTTCAC CTTCAGACCC CTCCACCCG CACCTCCGCC 660
 TCCTCATGCC TGACCTGTG CCAGGAAGCC ACCCCTGCA GCGGACTCTC TTCAGAGGAG 720
 ATCAATGACT ACCCGCAGCC AGCCAGCCC AGCTGCTCCA GCTCCCCCAA CCAGCAGCGA 780
 GGATPCAGTC CATCTGCATA ACAGCTGGGT CCTGAACAGC AACATACCAT TGGAGACCAG 840
 GCATTCCTCTG TTCAACCATG GATCTGGTTC CTCTGCGATC TTCACTGAGC CCAGTCAGAA 900
 CTACCTCTCTG ACATCCAATA CCGTGTACTC GCCCCCTCCC AGGCTCTTTC CTCGAAGCAC 960
 CTTTTCCTGA CCTGCCTTTA CCTTAACAA ACCTTACAGG TGCTGCAACT GGAAGTGCAC 1020
 AGCATTTAGC GCCACTGCAA TCACAGTGAC TTTGGCCTTG TTAATAGCCT ATGTGATTGC 1080
 AGTGCAATTG TCTGGCCTGA CTTGGCAGTT GCAACCAAGT GAAGGAGAGC TGTATGCAAA 1140
 TGGAGTTAGC AAAGGGAAAC GGGGGACCGA GTCCATGGAC ACTACTTACT CTCCAATTGG 1200
 AGGAAAAGTT TCTGATAAAT CAGAGAAAA AGTGTTCAG AAGGACGGG CGATAGACAC 1260
 TGGAGAAAGT GACATTTGTT GACAGGTCAT GCAGACCATT CCACCTGGTT TATTCTGGCG 1320
 TTTCCAGATT ACTATCCACC ATCCAATATA TCTGAAGTTC AATATTTCCT TAGCCAAGGA 1380
 CTCTCTGCTG GGAATTTATG GCAGAAGAAA CATTCACCT ACACATACTC AGTTTGATT 1440
 TGTAATAACT ATGATGGCA AACAGCTGGT CAAGCAGGAC TCCAAGGGCT CTGATGATAC 1500
 ACAGCACCTC CCTCGGAACC TGATCTTAAC TTCGCTTCAG GAGACAGGTT TCATAGAGTA 1560
 TATGGATCAA GGACCTTGGT ATCTGGCGTT TTACAATGAT GGAAGAAAGA TGGAGCAAGT 1620
 ATTCGTGTTA ACTACGCAA TTGAAATAAT GGATGACTGT TCAACCAATT GCAATGGAAA 1680
 TGGAGAGTGT ATCTCTGGCC ATTGTCAATT TTTCCAGGA TTCCTTGAC CTGACTGTGC 1740
 TAGAGATTCC TGCCCTGTGC TGTGTGTGG GAATGGAGAA TACGAGAAAG GACACTGTGT 1800
 CTGCCGGCAT GGCTGGAAGG GGCAGAGTGT TGACGTTCCG GAAGAACAA GCATTGATCC 1860
 AACATGCTTT GGCACGCGCA CTTGCATCAT GGGAGTCTGC ATCTGTGTGC CAGGATACAA 1920

	AGGAGAAATA	TGCGAGGAAG	AGGACTGCCT	AGACCCAATG	TGTTCCAACC	ATGGCATCTG	1980
	TGTAAAAGGA	GAATGTCAC	GTCTACTG	CTGGGGAGGA	GTAACTGTG	AAACACCACT	2040
	TCCTGTATGT	CAAGAGCAGT	GCTCAGGACA	CGGAACCTTT	CTTCTGGACG	CTGGAGTATG	2100
	CAGCTGTGAT	CCCAAGTGGA	CAGGATCTGA	CTGCTCAACA	GAGCTGTGTA	CCATGGAGTG	2160
5	TGGTAGCCAT	GGAGTCTGCT	CAAGAGGAAT	TTGCCAGTGT	GAAGAAGGCT	GGGTAGGACC	2220
	AACATGTGAG	GAACGCTCCT	GTCAATCTCA	TTGTACTGAG	CATGGCCAAT	GCAAAGATGG	2280
	AAAATGTGAG	TGTAGCCCTG	GATGGGAGGG	CGACCACTGC	ACAAATGCTC	ACTACTTAGA	2340
	TGCTGTCCGA	GATGGCTGCC	CAGGGCTCTG	CTTTGGAAAT	GGACGATGTA	CCCTGGATCA	2400
10	AAATGGTTGG	CACGTGTGTG	GTCAGGTGGG	TTGGAGTGGG	ACAGGCTGCA	ATGTTGTTCAT	2460
	GGAAATGCTT	TGTGGAGATA	ACTTGGACAA	TGATGGAGAT	GGTTTAACCG	ACTGTGTGGA	2520
	TCCTGACTGT	TGTCAACAAA	GCAACTGTGA	TATAAGTCTC	CTCTGCCAGG	GCTCACCAGA	2580
	TCCTCTGTAC	CTCATTGAGC	AAAGCCAAAC	TCTCTTCTCT	CAGCACACTT	CAAGACTTTT	2640
	TTATGATCGA	ATCAAAATTC	TCATTGGCAA	GGACAGTACT	CATGTTCATC	CTCCTGAGGT	2700
15	GTCAATTTGAC	AGCAGGCGGT	CCTGTGTGAT	TCGAGGCCAA	GTGGTGGCCA	TAGATGGAAC	2760
	TCCTCTAGTG	GGAGTGAATG	TCAGTTTCTT	GCACCACAGT	GATTATGGGT	TTACCATCAG	2820
	CCGGCAAGAT	GGAAAGCTTG	ACCTCGTGGC	CATCGGTGGC	ATCTCTGTCA	TCTTAATCTT	2880
	CGACCGATCC	CTTTCTCTGC	CTGAGAAGAG	AACACTCTGG	TTGCCCTTGA	ATCAGTTTAT	2940
	TGTGGTAGAG	AAAGTCACCA	TGCAGAGAGT	TGTATCAGAC	CCGCCATCCT	GCGATATCTC	3000
20	CAACTTTATC	AGCCCAAAAC	CTATTGTGCT	TCCTTCACCG	CTCACATCAT	TTGGAGGGTC	3060
	CTGTCCAGAG	AGGGGAACAT	TTGTTCTCTG	GCTGCAGGTT	GTACAGGAGG	AAATTCACAT	3120
	TCCTCTCCAG	TTTGTGAGGC	TGAGTTACCT	GAGCAGCCGC	ACCCCTGGGT	ATAAAACCTT	3180
	GCTACGGATC	CTTCTGACAC	ATTCAACGAT	TCCCCTAGGC	ATGATAAAAG	TACACCTCAC	3240
	AGTAGCTGTG	GAAGGGCGAC	TCACACAGAA	GTGGTTTCCC	GCCGCAATTA	ATCTTGTCTA	3300
25	CACATTTGCT	TGGAACAAGA	CCGATATCTA	TGGACAGAAG	GTTTGGGGCC	TGGCAGAGGC	3360
	TTTGGTATCT	GTGGGATATG	AATATGAAAC	GTGCCCTGAC	TTTATCTCTT	GGGAGCAAAG	3420
	GACAGTCGTT	TTACAAGGTT	TTGAGATGGA	TGCTTCTAAC	CTAGGAGACT	GGTCTTTGAA	3480
	TAAACATCAC	ATTTTGAATC	CTCAAGTGG	AATCATACAT	AAAGGGAATG	GAGAAAAATAT	3540
	GTTCATTTCC	CAGCAGCCCC	CAGTCATATC	AACCAATAAT	GGTAATGGAC	ACCAAAGGAG	3600
30	TGTAGCTGCT	ACCAACTGCA	ATGGCCGAGC	CCACAACAAC	AAACTCTTTG	CTCCTGTGCG	3660
	CTTAGCTTCT	GGCCCTGATG	GCAGTGTGTA	TGTTGGCGAC	TTCAATTTTG	TAAGGAGAAAT	3720
	ATTTCCCTCG	GGAAACTCCG	TTAGTATTTT	GGAAATTAAG	ACAAGTCTCG	CTCACAATAA	3780
	CTATCTGGCT	ATGGACCTTG	TGCTGAAATC	ACTCTATCTA	TCAGACACCA	ATACTCGCAA	3840
	AGTCTACAAG	TTGAAATCTC	TTGTGGAGAC	GAAGAATCTG	TCCAAGAATT	TTGAAGTGGT	3900
35	GGCAGGAAT	GGTGAATCAG	GCCTTCCCCT	TGACCAGAGT	CATTGTGGAG	ATGGTGGGAG	3960
	AGCATCGGAA	GCTTCACTGA	ATAGCCCTCG	AGGCATCACA	GTGTGATAGG	ATGGATTTAT	4020
	TTACTTTGTG	GATGGGACTA	TGATTCGCAA	AATTGATGAG	AATGCTGTGA	TCACAACCTG	4080
	AATCGGCTCA	AATGGTCTGA	CTTCCACACA	ACCACCTGAG	TGTGACTCAG	GAATGGACAT	4140
	CACCTAGGTT	CGATTAGAGT	GGCCAACAGA	CCTTGCAGTA	AATCTATGGA	ACAATTCATT	4200
40	GTATGTCTTG	GATAACAACA	TTGTGCTGCA	AATTTCTGAG	AACAGGCGTG	TTCCGATCAT	4260
	CCGAGGACCG	GCCATTCACAT	GCCAGGTGCC	AGGCATCGAT	CATTTCCCTG	TCAGCAAGGT	4320
	AGCAATTCAC	TCCACTCTAG	AGTCAGCGAG	GGCCATCAGT	GTCTCCCAAC	GCGGGCTGCT	4380
	CTTCACTAGCT	GAACACAGAG	AGAGGAAAGT	AAACCGCATT	CAGCAAGTAA	CCACCAATGG	4440
	GGAGATCTAC	ATCATCGCTG	GTGCCCCAC	TGACTGTGAC	TGCAAAATTG	ATCCAAACTG	4500
45	TGACTGTTTT	TCAGGTGATG	GTGGCTATGC	CAAAGATGCA	AAGATGAAAG	CCCCCTCCCT	4560
	CTTAGCAGTG	TCGCCCTGAT	GAACCTCTTA	TGTGGCAGAC	CTCGGAAATG	TTCGAATTCT	4620
	TACCATCAGC	AGGAACCAAG	CCCACCTGAA	TGACATGAAC	ATTTATGAGA	TTGCTTCACC	4680
	CGCTGATCAG	GAACGTGAC	AGTTCACTGT	AAATGGAACC	CACCTACACA	CCCTGAACCT	4740
	GATAACAAGG	GACTATGTTT	ATAACTTCAC	CTACAATTCT	GAAGTGAAGT	TGGGCGCGAT	4800
50	TACCAGCAGC	AATGGCAATT	CAGTGCACAT	TCGCCGTGAT	GCAGGCGGAA	TGCCGCTATG	4860
	GCTTGTGGTG	CCTGGCGGAC	AAGTATACTG	GCTGACTATA	AGCAGCAATG	GAGTCTTGAA	4920
	AAGAGTGTCA	GCCCAAGGCT	ATAATCCGGC	CTTAATGACC	TATCCAGGAA	ACACAGGGCT	4980
	TCTGGCTACC	AAAAGTAACG	AAAATGGATG	GACAACCGTT	TATGAGTATG	ACCCGAGGGG	5040
	ACACCTGACC	AATGCAACGT	TTCCCCTGAG	AGAGGTCAGC	AGCTTCCACA	GTGACCTGGA	5100
55	GAGCTGACCA	AAAGTGGAGC	TAGATACTTC	CAACCGTGAA	AATGTCTCTA	TGTCAACCAA	5160
	CTTGACGGCA	ACTAGTACCA	TATATATTTT	AAAACAAGAA	AATACTCAA	GTACCTATCG	5220
	GGTGAATCCA	GATGGTTCCC	TGCGTGTGAC	TTTTGCCAGC	GGGATGGAGA	TGGGCTCAG	5280
	CTCAGAGCCC	CACATCCTGG	CAGGGGCGAG	CAACCCTACC	CTGGGCAAT	GCAACATCTC	5340
	ATTGCCCGGA	GAGCACAATG	CAAACTCAT	CGAGTGGCGG	CAGAGGAAGG	AGCAAAACAA	5400
60	AGGCAATGTT	TCGGCTTTTG	AAAGGAGGCT	GAGGGCCAC	AACAGAAACC	TACTCTCCAT	5460
	AGATTTTGAT	CATATAACCC	GCACAGGAAA	GATCTATGAT	GACCATCGAA	AATTCACCCT	5520
	TCGAATTTCT	TATGACCAGA	CTGGGCGACC	CATTCTGTGG	TCTCCTGTAA	GCAGATATAA	5580
	TGAAGTGAAC	ATCACATATT	CACCTTCGGG	ATTGGTGACG	TTTATTCAAA	GAGGAACGTG	5640
	GAATGAAAAA	ATGGAATATG	ACCAGAGTGG	GAAAAATTAT	TCAAGAACTT	GGGCTGATGG	5700
65	GAAAAATTGG	AGCTATACCT	ACTTAGAAAA	ATCTGTGATG	CTTCTCTTAC	ACAGCCAGCG	5760
	GCGTTACATC	TTTGAGTATG	ACCAATCAGA	TTGCCCTGCT	TCAGTTACCA	TGCCCTAGCAT	5820
	GGTGGCGCAC	AGCTTACAAA	CCATGCTTTC	AGTGGGCTAC	TACCGTAATA	TCTACACCCC	5880
	ACCGGACAGT	AGCATTCTTT	TTATCCAAGA	CTATAGTCGA	GATGGCCGAT	TGCTACAGAG	5940
	CCTGCATCTG	GGGACAGGCT	GCAGAGTCTT	ATACAAGTAC	ACCAAGCAAG	CAAGGCTTTC	6000
70	TGAGGTTCTC	TATGATACCA	CTCAGGTCAC	ATTAACATAT	GAAGAGTCTT	CTGGAGTGAT	6060
	TAAACAATAA	CACCTGATGC	ATGACGGATT	CATCTGCACA	ATCAGATACA	GGCAACACAG	6120
	ACCTCTTATT	GGACGCCAGA	TTTTCAGATT	CAGTGAAGAA	GGCCTTGTGA	ATGCACGGTT	6180
	CGACTACAGC	TACAACAATT	TCCGAGTCAC	AAGCATGCAA	GCTGTAATCA	ATGAACCCCC	6240
	TTTGCCTATA	GATCTTTTACC	GATATGTTGA	TGCTCTGTGC	AGAACAGAGC	AGTTTGGAAA	6300
75	ATTCAGTGTA	ATTAAATTAG	ATTAAATCA	GGTCATAACT	ACTACAGTGA	TGAAACACAG	6360
	CAAAATCTTC	AGTGCCAATG	GACAAGTCAT	TGAAGTCCAA	TATGAAATCC	TAAAGGCAAT	6420
	TGCTTACTGG	ATGACCATTC	AATATGATAA	TGTGGGCCGA	CATGGTAATA	TGTGCATAAG	6480
	GGTAGGAGTA	GATGCCAATA	TAACAAGGTA	CTTCTATGAA	TACGATGCTG	ATGGGCAACT	6540
	TCAGACTGTT	TCTGTAAATG	ACAAAACCCA	GTGGCGTTAT	AGTTACGATC	TGAATGGAGA	6600
80	CATCAACCTC	TTAAGCCATG	GGAAGAGTGC	TCGCTTACT	CCTCTCCGAT	ATGACCTCCG	6660
	AGACCGCATC	ACCAGATTAG	GAGAAATCCA	GTATAAAATG	GATGAAGATG	GCTTTCTGAG	6720

	GCAGAGGGGA	AAATGATATTT	TTGAATATAA	TTCTAATGGC	CTGCTGCAGA	AAGCCTACAA	6780
	TAAGGCTTCT	GGCTGGACTG	TGCAGTATTA	CTATGATGGG	CTTGGGCGAC	GTGTGCGGAG	6840
5	TAAGTCCAGC	CTAGGGCAGC	ACCTTCAGTT	CTTTGTCGAC	GCGACCGCGA	ACCCCATAG	6900
	AGTTACTCAT	TTGTACAACC	ACACAAGCTC	GGAGATTACA	TCTCTGTATT	ATGATCTCCA	6960
	AGGTACACCT	ATTGCCATGG	AGTTAAGCAG	TGGTGAAGAA	TATTTATGTAG	CCTGTGATAA	7020
	TACAGGTACC	CCACTAGCTG	TGTTACGACG	CCGAGGTCAG	GTCAATAAAGG	AGATACTATA	7080
	CACACCTTAT	GGCGATATCT	ATCATGACAC	TTACCCTGAC	TTTCAGGTCA	TAATTGGTTT	7140
	TCATGGAGGA	CTCTATGATT	TCCTTACTAA	ATTAGTGCAC	CTGGGGCAAA	GGGATTATGA	7200
10	TGTTGTTGCT	GGCAGATGGA	CAACGGCCTA	TCATCACATA	TGGAACACAGT	TGAACCTCCT	7260
	TCCTAAACCA	TTCAACCTCT	ACTCCTTTGA	AAATAACTAC	CCAGTTGGCA	AAATTCAGA	7320
	TGTTGCCAAG	TATACCACAG	ACATCAGAAG	TTGGTTGGAG	CTATTTGGTT	TCCAATTACA	7380
	CAATGTACTA	CCTGGATTTT	CCAAACCTGA	ATTAGAAAAT	TTAGAATTAA	CTTACGAGCT	7440
	TTACGCGCTT	CAGACAAAAA	CTCAAGAGTG	GGATCCTGGA	AAGACTATCC	TGGGCATTCA	7500
15	GTGTGAACCT	CAGAAAACAGC	TCAGGAATTT	CATTTCTCTG	GACCAACTAC	CTATGACTCC	7560
	CCGATACAAT	GATGGACGGT	GCCTTGAAGG	AGGGAAGCAA	CCAAGGTTTG	CTGCTGTCCC	7620
	TTCTGTTTTT	GGGAAAGGTA	TAAAAATTTG	CATCAAGGAT	GGCATAGTAA	CAGCTGATAT	7680
	TATAGGAGTA	GCCCAATGAAG	ATAGCAGGCG	GCTTGCTGCC	ATTTCTCAATA	ATGCCCATTA	7740
	CTTGAAAAAC	CTACATTTTA	CCATAGAGGG	GAGGGACACT	CACTACTTCA	TTAAGCTTGG	7800
20	GTCTCTGGAG	GAAGACCTGG	TGCTCATCGG	TAACACTGGG	GGGAGGCGGA	TTCTGGAGAA	7860
	TGGTGTCAT	GTACATGTGT	CCAGATGAC	TTCTCTGTGT	AATGGGAGGA	CTAGACGGTT	7920
	TGCAGATATT	CAGCTCCAGC	ATGGAGCCCT	GTGCTTCAAC	ATCCGGTATG	GGACAACCTGT	7980
	CGAAGAGGAA	AGAATTCACG	TGTTGGAGAT	TGCCAGACAG	CGCGCAGTGG	CCCAGGCTGG	8040
	GACTAAGGAA	CAAAGAAGGC	TGCAAGAGGG	GGAAGAGGGG	ATTAGGGCAT	GGACAGAAGG	8100
25	GGAAAAGCAG	CAGCTTTTGA	GCACTGGGCG	GGTACAAGGT	TACGATGGGT	ATTTTGTTTT	8160
	GTCTGTGTAG	CAGTATTTAG	AACCTTCTGA	CAGTGCCAAT	AATATTCAC	TTATGAGACA	8220
	GAGCGAAATA	GGCAGGAGGT	AACAAAAATA	TCTCTGCCCT	TGCGTCACCA	AAGACTGCCT	8280
	GTMTTAAAAA	CATAAAATGG	TTTATTTGTAT	TGGTTTTCTA	GATCAGAACT	CTGTATATGT	8340
	AAATATGGAG	GA AAAACATA	TCCAACCTGCC	TTTCAATGTG	ACGGAAGATG	GTATTTTAAT	8400
30	ATTGTTTGT	TAAACTCTTT	AAGAAATGAC	AGAGATTTTT	AGTTCTTGTG	TGGCAGTATT	8460
	CAAAAATAACA	CAAGTAGAAT	TCAAAACAGCT	AAAAACAGTT	TTAGAAAAGC	ACCACCTTCA	8520
	ATTTCGCGAG	CCATGCATAT	GTTCCAATAT	CCAGAAAAGAA	CCCAAGGTTT	TCTATCTCTA	8580
	TTGTGAGAAG	CAGTTTCATC	CTTAACTGTT	GGCAGAACTT	ACGGGCTATT	TGAATAGGTG	8640
	GTGCAATAGT	ATCTGAAACT	TGCTTTTCGA	AAGACTGCCA	GCCCTTTGAC	GTMTTCCAGA	8700
35	TCTGTTATAG	GAACCTTAA	AACAGGTGTA	AAATGTCTTC	AGCCACCATC	TCCTAGAGTG	8760
	AGGACCCAAT	TGCCCTTCTC	TCTTGATTAT	TCTCTCTTGC	TGTTTAAAGT	AAATGCCATA	8820
	TTGTTGTGCT	GTGTTTGTGC	GTGTGGTGGC	TGGGTTCTGT	CTACCATGCT	TCCCTGTGGG	8880
	TGTGGTAACC	AGACTGTATA	GCCGCTATTT	GCTCGTGTGT	ACATGATACC	AAAGCAGCTG	8940
40	GCCAGCGTGA	CTCTCTCAC	ACGACCTGTT	TTGACTCAAT	TTTTTACTAA	AAGTTGTTCA	9000
	GCTGTATTGG	TATCATGTAA	ACATAGCTTT	TATTAACCTG	GGTAGGAATT	TCTCATTTAT	9060
	ATATAGGATG	TGTTTGTGGT	ATAGTTTCAC	ATTAGTGATT	CAGTATCTAT	ACACTGACCC	9120
	AAATGGTTTG	TGCACATGAA	CGGTAATTTA	CTTAAAAGTA	TGATTTCTGGT	ACAAAAACAA	9180
	ACAAAGGCTT	TAGCAGGCAT	ACGTGTCTGG	GATGCCGATA	CATACATTAA	CTACTACTGC	9240
45	AGAAATTCAT	AAGAGCCAAA	ACCTTAAAAA	AATAGACCTG	GTACTTAAAT	GAAAGTACTA	9300
	AAGGGAAGAC	CAGACCAAAC	ATCACAGCAG	TTGCTGCCAC	ATTGTTTCAG	CCCCTTAGA	9360
	TTTATCTTTC	AAATGTACAA	TTCTGTATTG	AACATCTCCC	AGCCATCTTC	AGGAAATCGA	9420
	ATCAAGTAAA	TCTTTTCCAA	CCGAAAACAT	TTCAACTAAC	TATAGAGAGG	CAGACTCATT	9480
	TTTACTAAAA	TAATTTATAC	AGTTAGTTAT	TTTCGTTCTC	CGTACTTACC	CATTTATCTT	9540
50	TATTTAATCG	TCTCTACTGC	CTAGGAAAAT	AACATTTTTC	CAGGACGGGT	TATTTGTTC	9600
	GCGATCATTT	AAAATTTGGA	GAAAGGTCAG	GATTAGTGTT	AATATCAGCT	GCAGTTTCTC	9660
	AATCTCTAGG	AATCTGTCAG	TAAAACAAGC	CCCTTGGTGA	GCTGGAAGAT	TTGTGCCCTG	9720
	TGACAAAAGAG	ATAGTTTGTG	AAATGCTGTG	TAATTTGTAAG	TTACCACAAA	TGAAAAATACA	9780
	TGACAGCACA	ATGTGGCCCG	TAGAAAATTC	CCCTGAGCCA	GCTTCTGCAC	TTTCATCACC	9840
55	GAATCTGAAC	ATTTGCTATG	TCTGAAGGCA	AATTTATGAT	GGAATGTTAG	TTTGGATTCT	9900
	TTCCAGATGC	TACCTTAAAT	CAGTGTGGGG	TCATTGCTAT	GCTTTGCGAT	GACAGTTTCT	9960
	TTGAAAATAT	GCAAAGTCAT	AAGCTCATGT	TAAGGTTTTT	CAAGAGTCTG	CCTCCTACTA	10020
	CACAAAAGGAA	AGCAAGGGAA	AGGAATGAC	CCTGGCAAAC	AGTAGGGAAG	GGTGTATTCA	10080
	AACATTTTCAT	TTTCAAAACC	TTCCGGGTTAG	AATACCACCT	ACATATGTAT	TCTGAGAGAC	10140
	AGAATTCATG	AGGAACCTAT	CTCTCTTTAT	AACCTGGAAC	ACACCAGCTT	GATATATTGC	10200
60	TAATCCATAC	TAAAATCATA	TTATTGGGTT	TTTTCTGAAT	CAGGCCGTGA	TTAATGGTAC	10260
	AGTATTTTAT	CAGAAATGGA	TTCTAAAAAT	ACTAACAAAC	TTGTTGAAAA	TTTGAATACC	10320
	TCCACACCAA	CCTAAAATGT	GACCTTAAAT	TCTTAGAAC	TCTGATGTTT	TTTTAAATTA	10380
	ATGGAATAAT	AATTTGTGAA	CTGTATATAG	AGAGTGCATT	CATAAATGTG	ATTATGTATT	10440
	TTATCACAAA	TCCAAAATGT	CAATATTAGA	GTCTATTTTG	CTTATATTTT	AAGCAATTAT	10500
65	ACGTTTTTGC	AATTCATTGA	TGATGTATCA	TTTTCAAACT	GCTTTAAATA	TCCATTAGAA	10560
	ACAAATATTT	GAAGCTTTTA	CTTAATAGTG	ATTACCTTGA	ACTGTGCATT	TCTAGTTTGT	10620
	AATACGTATT	TGGTTGGTTT	GTGCTTTTAT	TTTGTAAAG	TTACATTTGT	ATTATATTCA	10680
	GGAAATGCAC	TTTTTTATTAC	TTACAGCTGT	GGTTTTAATA	CTGCTTGAA	CTATTATTAT	10740
	TCTTTTACAA	ACTCTTAAAG	CTTGAGGGAG	GAAAGAAAAA	AAAAACAAAA	CTACTAATCA	10800
70	GTAGTAAATC	GAAGAGAAAC	ATTTTGGCAT	TTCTTAAAGAA	GAAGATGGAG	ATATTGAGTA	10860
	TATCACTTCC	TATTCAGCTG	AATAGAAAGA	ATGCCCTTCT	TGACTTGCAG	TTCTGCAGTT	10920
	TAAATTTATG	AAAGAACAAT	TGTTTTCAT	TTCCCTGATG	AAGTAAAGC	ATTTTTCAGA	10980
	GAAACATATG	AATTTCTCAT	ACCCAGCAGA	CAGATGGCTG	ACACTGCACA	GCCACACACC	11040
	ATTCGAGTAA	GTTAAAGTGA	GAGCATAGTA	GTTGGACTCT	CCTATGAAGA	ACATTTCTGGG	11100
75	CTGGAGGCAG	GGAACTACTC	ATGGTTGTTT	CTTTTCTCTA	CTTAAGCCCA	TTTTTGTGT	11160
	GCTTTTCTGT	TTTGTTTTGT	TTTCACTCTT	GCACCTACAGT	CTAGAGATCC	AAATGAACTG	11220
	AAAAGTTCAA	AGTTTAAAC	ATTTTAAATAT	GTTTACTTTT	AGTTGTCTAT	CTAATCTGTA	11280
	TTGATTAGAA	CCATGACTCC	TGAAGGAAAG	GGAAATAAAT	CTCAATTCAT	ACTAACTTGC	11340
	AACAAAACAC	TTTTTACCATA	TAAATAAGTA	TATGATTTAT	TTTTAACCCA	AAAAATGTAT	11400
80	AAAATAAGTG	TGTCCTTTAT	TGTCATTTTA	TCGAGAAGAT	CTATAATATA	TAGACTACAT	11460
	ATATATAATA	TATACAACAT	AGCCAAATGT	ATGAAAACCT	GACAAATGAT	AATTTGGAAT	11520

TCACATGCTA CCTATGTAGA CAGGTATGAA ATTAAGTTAT AATTTTCATG AGACATTTTC 11580
 ATCACTGTTG ACACAGTTTC AAGGCATTCC ATCATGTTAT TTTGACTCTT TTTCTTTTTT 11640
 TTTTCTTTAA AAATATATTT TTAAGTAGAC CAGGCCCCAC TATAATATCA CTTAAGAGAG 11700
 TCAGGGCAAA GTTTTTCGAT TTATGAAGAT GTGTTTCATGT AAGGGTGATT GTAATGGAGT 11760
 TCATTGGTAA TAGAAGCAAA AGTACAGTAA CGAAGTATTG AAAAGAAAAT TTTGGAGACA 11820
 TGGAGACATA TTATATATAG CTGTGGGAAA GACATAAGGC TACAGATGGA ATGGAACATT 11880
 CCTGTTTTCT TGAGAAATTT CACATACACA TAGCTGACCT GACTAGTACT TCAGCTCTTC 11940
 CACAGCCTTC TATAAAGGTT CTTTCTTCTG CAAAGAAAAC AAAACAAAAC AAAACAAAAC 12000
 AAAAAAAAAC AAAAAAAGCG CAAAAAACAA AAAAAACAAA AAAAGCAAAG TAAATTTTAA 12060
 AAATACAGAA AACAAACAAC AAAAAAGAA TCAACCATAA ATAGTGACTA TTATTTTCAG 12120
 TGTGTCCTTC ATGTGAAAGC TATTAAGGAC CAAATATACT ACTGTTTCATA AGAAGAAATT 12180
 ACTTCTTAAA CAGTAACTGA AAATACTTAG AGTTAAACTT GCTGTGGATT TTGCTCTTGC 12240
 AGTTGTCATC TTACATTATT TGTCAAAGGA AATGTGTTTG GCAGTTAAAA ATCTTTTCCT 12300
 AGATTTAGTG GTGGACTTTA ACCTCTTAAA TAAATGTTAG TATATCAGAT TGTGTCCTTG 12360
 AAAAAATTTT TACTTGTATG AATCATGACA ACGTCTAAAT CTTTACTATT CTCTGCGCAA 12420
 AAGCATCAGT AGGAAAGAG GCGAAAAGA GAAGTATAGC CTTTATGTCA GAAAAACATT 12480
 CTTTTTCAGT GCTTACTTTC TCATGAAAAG TAAAGATGTT TACAGTGAT GCCAAGTTTT 12540
 CAGTTTCTGT ATAACAACAG GTAGAGGTTT TAATCATATT GAAAAATGTT TTATAATGGT 12600
 CTGAGCCATG TTGCTAGGAA ACAAAGGTT CCAATTTTGT ATTCTGCTC TCCTGTGCTG 12660
 AAAAGTGACT GGATACGTGA CAGGTTTCATG TTCTCTGGCT GCAGTTAAAT GGTCTTTTGC 12720
 ATTTTGCTCT GGCTTTCAGG CCAGAAGCAT GCATTTTCTT ACAAGAGCAT CACAACAACA 12780
 TGCTGTAAAT ATTTAAAGTT AAACATTATG TGTGATATT TGAAAGAAA GTACTTTGAA 12840
 TATTTTCATT TTAATAAATA AAATTGCCAA TGAAAAAAA

SEQ ID NO:230 PEZ2 Protein sequence:

Protein Accession #: NP_055068

1 11 21 31 41 51
 MEQTDCKPYQ PLPKVKHEMD LAYTSSSDES EDGRKPRQSY NSRETLHEYN QELRMVNSQ 60
 SRKRKEVEKS TOEMFECETS HTLCSTGYQTD MHSVSRHGYY LEMGSDVDTE TEGAASPDHA 120
 LRMRWIRGMS EHSSCLSSRA NSALSLTDDP HERKSDGENG FKFSFVCCDM BAQAGSTQDV 180
 QSSPHNQFTF RLPFPFPFPF HACTCARKPP PAADSLQRRS MTRRSQPSFA APAPPTSTQD 240
 SVHLHNSWVL NSNIPLETRH SLFKHSGSGS AIFSAASQNY PLTSNTVYSP PPRPLPRSTF 300
 SRPAFTFNKP YRCCNWKCTA LSATAITVTI ALLLAYVIAV HLFGLTWQLQ PVEGELYANG 360
 VSKGNRGTES MDTTSPPIGG KVSDEKSEKV FQKGRADITG EVDIGAQVMQ TIPPGLFWRP 420
 QPTIHHPIYL KFNISLAKDS LLGIYGRRI PPTHQDFDV KLMDGKQLVK QDSKGSDDTQ 480
 HSPRNLILTS LQETGFIEYM DQGPWYLAFF NDGKKMEQVF VLTATAIEIMD DCSTNCGNG 540
 ECISGHCHCF PGFLGPDCAR DSCFPLCGGN GEYKGGHCV C RHGWKGPEDC VPPEQCIDPT 600
 CFHGHGTCIMG VCICVPGYKG EICEBEDCLD PMCSNHGICV KGECHCSTGW GGVNCEPFLP 660
 VQEQCSGSHG TFLDLGAVCS CDPKWTGSDC STELCTMECG SHGVCSRGIC QCEEGWVGPT 720
 CEERSCHSHC TEHGQCKDGK CECSPGWEGD HCTIAHYLDA VRDGCFLGCF GNGRCTLQDN 780
 GWHCVQVQGW SGTGCVNVME MLCGDNLDND GDGLTDCVDP DCCQSQSNYI SPLCQGSDDP 840
 LDLIQQSQTL FSQHTSRLFY DRIKFLIGKD STHVIPPEVS FDSRRACVIR GQVVAIDGTP 900
 LVGVNVSLFL HSDYGFITSR QDGSFDLVAI GGISVILIFD RSPFLPEKRT LWLFWNQFIV 960
 VEKVTMQRVV SDPPSCDISN FISPNIPIVLP SPLTSFGGSC PERGTIVPEL QVVQEEIPI 1020
 SSFVRLSYLS SRTPGYKTL RLLTHSTIP VGMIVHLLTV AVEGRLLTQW FPAAINLYVT 1080
 FAWNKTDIYQ QKVGALAEAL VSVGYEYETC PDFILWEQRT VVLQGFEMDA SNLGDWSLNL 1140
 HHILNFPQSGI IHKGNGENMF ISQPPVIST IMGNGHQRSV ACTNCGPAH NNLKLFAPVAL 1200
 ASGPDGSVYV GDFNFVRRIF PSGNSVSILE LSTSPAHKYY LAMPVSESL YLSDTNTKRV 1260
 YKLKSLVETK DLSKNFEVFA GTGDQCLPFD QSHCGDGGRA SEASLNSPRG ITVDRHGFIV 1320
 FVDGTMIRKI DENAVITTVI GSNGLTSTQP LSCDSGMDIT QVRLEWPTDL AVNPMDNSLY 1380
 VLDNNIVLQI SENRRVRIIA GRPIHCQVPG IDHFLVSKVA IHSTLESARA ISVSHSGLLF 1440
 LAETDERKVN RIQOVTTNGE IYIIAGAPTD CDCKIDPNCD CFSGDGGYAK DAKMKAPSSL 1500
 AVSPDGTLYV ADLGNVRIIT ISRNQAHLD MNIYELASPA DQELYQFTVN GTHLHTLNL 1560
 TRDYVYNFTY NSEGLDAIT SSNGNSVHIR RDAGGMPLWL VVPGQVYWL TISSNGVLKR 1620
 VSAQGYNPAL MTYPNGTGLL ATKSNEGWT TVYEYDPEGH LTNATFPTGE VSSFHSDLEK 1680
 LTKVELDTSN RENVMSTNL TATSTIYILK QENTQSTYRV NPDGSLRVTF ASGMEIGLSS 1740
 EPHILAGAVN PTLGKCNISL PGEHNANLIE WRQRKEQNK G NVSAFERRLR AHNRLNLSID 1800
 FDHITRTGKI YDHRKFTLR ILYDQTGRPI LWSFVSRYNE VNITYSPSGL VTFIQRGTWN 1860
 EKMEYDQSGK IISRTWADGK IWSYTYLEKS VMLLHQSRR YIFEYDQSDC LLSVTMPMSV 1920
 RHSLQTMLSV GYRNIYTPP DSSTFSFIQDY SRDGRLLQTL HLGTRRRVLY KYTKQARLSE 1980
 VLYDPTQVTL TYEBSGVIK TIHLMDHGF CTIRYRQTGP LIGRQIFRS EELVNFARFD 2040
 YSVNFRVTS MQAVINETPL PIDLYRYVDV SGRTEQFGKF SVINYDLNQV ITTVMKHTK 2100
 IFSANGQVIE VQYEILKAIA YWMTIQYDNV GRHGNMCIKV GVDANITRYF YEYDADGQLQ 2160
 TVSVNDKTQW RYSYDLNGDI NLLSHGKSAR LTPLYDLRD RITRLGEIQY KMDGDFLRQ 2220
 RGNDIFEYNS NGLLQKAYNK ASGWTVQYYY DGLGRRVASK SSLGQHLQFF VDANANPIRV 2280
 THLYNHTSSE ITSLEYDLQG HLIAMELSSG EBYVACDNT GTPLAVFSSR GQVIKEILYT 2340
 PYGDIYHDTY PDFQVIGFH VHLGQRDYDV VAGRWITAYH HIWKQLNLLP 2400
 KPFNLYSFEN NYFVGKIQDV AKYTTDIRSW LELFGFQLHN VLPGFPPKPEL ENLELTIELL 2460
 RLQTKTQEWDP PKGTTLIGIQ ELQKQLRNF SLDQLFMTPR YNDGRCLGEG KQPRFAAVPS 2520
 VFGKGIKFAI KGIVITADI GVANEDSRRL AAILNNAHYL ENLHFTIEGR DTHYFIKLS 2580
 LEEDLVLIGN TGGRRILENG VNVTVSQMTS LLNGRTRRFA DIQLQHGALC FNIRYGTIVE 2640
 EEKNHVLEIA RQRAVAQAWT KEQRRLOQEGE EGIRAWTEGE KQQLLSTGRV QGYDGYFVLS 2700
 VEQYLELSDS ANNIHFMRQS EIGRR

SEQ ID NO:231 PFD4 DNA SEQUENCE:

Nucleic Acid Accession #: NM_000441

Coding sequence:

225-2567 (underlined sequences correspond to start and stop codons)

	1	11	21	31	41	51	
5	CTCAGCCTTC	CCGGTTCGGG	AAAGGGGAAG	AATGCAGGAG	GGGTAGGAT	TCTTTCTCTGA	60
	TAGGATCCGT	TGGGAAAGAC	CGCAGCCTGT	GTGTGTCTTT	CCCTTCGACC	AAGGTGTCTG	120
	TTGCTCCGTA	AATAAAACGT	CCCACCTGCC	TCTGAGAGCG	CTATAAAGGC	AGCGGAAGGG	180
10	TAGTCCCGCG	GGCATTCCCG	GCGGGGCGCG	AGCAGAGACA	GGTCATGGCA	GCGCCAGGCG	240
	GCAGGTCCGA	GCCGCCGCG	CTCCCCGAGT	ACAGCTGCAG	CTACATGGTG	TCCGCCGCCG	300
	TCACAGCGA	GCTCGCTTTC	CAGCAACAGC	ACGAGCGCGC	CCTGCAGGAG	CGCAAGACGC	360
	TGCGGGAGAG	CCTGGCCAAG	TGCTGCAGTT	GTTCAGAAA	GAGAGCCTTT	GGTGTGCTAA	420
	AGACTCTTGT	GCCCATCTTG	GAGTGGCTCC	CCAAATACCG	AGTCAAGGAA	TGGCTGCTTA	480
15	GTGACGTCAT	TTCGGGAGTT	AGTACTGGGC	TAGTGGCCAC	GCTGCAAGGG	ATGGCATATG	540
	CCCTACTAGC	TGCAGTTCCT	GTCCGGATATG	GTCTCTACTC	TGCTTTTTC	CCTATCTCTGA	600
	CATACTTTAT	CTTTGGAAACA	TCAAGACATA	TCTCAGTTGG	ACCTTTTCCA	GTGGTGAGTT	660
	TAATGGTGGG	ATCTGTGTGT	CTGAGCATGG	CCCCCGACGA	ACACTTCTCT	GTATCCAGCA	720
	GCAATGGAAC	TGTATTAAAT	ACTACTATGA	TAGACACTGC	AGCTAGAGAT	ACAGCTAGAG	780
20	TCTGTATTGC	CAGTGCCTCG	ACTCTGCTGG	TTGGAATTAT	ACAGTTGATA	TTTGGTGGCT	840
	TGCAGATTGG	ATTCATAGTG	AGGTACTTGG	CAGATCCTTT	GGTTGGTGGC	TTCAACAACAG	900
	CTGCTGCCCT	CCAAGTGCTG	GTCTCACAGC	TAAAGATTGT	CCTCAATGTT	TCAACCAAAA	960
	ACTACAATGG	AGTTCCTCT	ATTATCTATA	CGCTGGTTGA	GATTTTTCAA	AATATTGGTG	1020
	ATACCAATCT	TGCTGATTTT	ACTGCTGGAT	TGCTCACCAT	TGTCGTCTGT	ATGGCAGTTA	1080
25	AGGAATTAAA	TGATCGGTTT	AGACACAAAA	TCCAGTCCC	TATTCCTATA	GAAAGTAATTG	1140
	TGACGATAAT	TGCTACTGCC	ATTTCATATG	GAGCCAACT	GGAAAAAAT	TACATGCTG	1200
	GCATTGTAA	ATCCATCCCA	AGGGGGTTTT	TGCCCTCTGA	ACTTCCACCT	GTGAGCTTGT	1260
	TCTCGGAGAT	GCTGGCTGCA	TCATTTTCCA	TCGCTGTGGT	GGCTTATGCT	ATTGCAAGTT	1320
	CAGTAGGAAA	AGTATATGCC	ACCAAGTATG	ATTACACCAT	CGATGGGAAC	CAGGAATTCA	1380
30	TTGCCCTTGG	GATCAGCAAC	ATCTTCTCAG	GATTCTTCTC	TTGTTTTGTG	GCCACCACTG	1440
	CTCTTTCCCG	CACGGCCGTC	CAGGAGAGCA	CTGGAGGAAA	GACACAGGTT	GCTGGCATCA	1500
	TCTCTGCTGC	GATTGTGATG	ATCGCCATTC	TTGCCCTGGG	GAAGCTTCTG	GAACCCCTGC	1560
	AGAAGTCGGT	CTTGGCAGCT	GTGTGAATTG	CCAACCTGAA	AGGGATGTTT	ATGCAGCTGT	1620
	GTGACATTCC	TGCTCTGTGG	AGACAGAATA	AGATTGATGC	TGTTATCTGG	GTGTTTACGT	1680
35	GTATAGTGTC	CATCAATTCT	GGGCTGGATC	TCGGTTTACT	AGCTGGCCTT	ATATTGGGAC	1740
	TGTTGACTGT	GGTCTGAGA	GTTCAAGTTT	CTTCTTGGAA	TGGCCTTGGA	AGCATCCCTA	1800
	GCACAGATAT	CTACAAAAGT	ACCAAGAATT	ACAAAAACAT	TGAAGAACCT	CAAGGAGTGA	1860
	AGATTCTTAG	ATTTTCCAGT	CTTATTTTCT	ATGGCAATGT	CGATGGTTT	AAAAAATGTA	1920
	TCAAGTCCAC	AGTTGGATT	GATGCCATTA	GAGTATATA	TAAGAGGCTG	AAAGCGCTGA	1980
40	GGAAAATACA	GAACAATA	AAAAGTGGAC	AATTAAGAGC	AACAAGAAAT	GGCATCATAA	2040
	GTGATGCTGT	TGCAACAAAT	AATGCTTTTG	AGCCTGATGA	GGATATTGAA	GATCTGGAGG	2100
	AACTTGATAT	CCCAACCAAG	GAAATAGAGA	TTCAGTGGAA	TGGGAACCT	GAGCTTCCAG	2160
	TCAAAGTGAA	CTTGTCCAAA	GTGCCAATCC	ATAGCCTTGT	GCTTGACTGT	GGAGCTATAT	2220
	CTTCTCTGGA	CGTTGTGTGA	GTGAGATCAC	TGCGGGTGAT	TGTCAAAGAA	TTCCAAGAA	2280
45	TTGATGTGAA	TGTGTATTTT	GCATCACTTC	AAGATTATGT	GATAGAAAAG	CTGGAGCAAT	2340
	GCGGGTTCCT	TGACGACAAC	ATTAGAAAAG	ACACATTCTT	TTTGACGGTC	CATGATGCTA	2400
	TACTCTATCT	ACAGAACCA	GTGAATCTC	AAGAGGGTCA	AGGTTCCTAT	TTAGAAAACA	2460
	TCACTCTCAT	TCAAGGATGT	AAAGATACCC	TTGAATTAAT	AGAAACAGAG	CTGACGGAAG	2520
	AAGAAGTTGA	TGTCCAGGAT	GAGGCTATGC	GTACACTTGC	ATCCCTGAAG	TGGGTTCCGG	2580
50	AGGTCTCTAT	GAGCAAGGAA	TACAAGACAA	AACCTTCTCA	ATGCATTGAC	TATTTCTTCA	2640
	GACTCAAAAC	ACTCATCTCT	TTTTCTATTA	AGCCATTGAA	AGAGAAGCAC	TAAGACTGCT	2700
	TCTAGGCTTT	ATTTATAAAA	TAAACACCTT	ATCCCTAACA	TGGGCAAAAT	GGCTAGAAAT	2760
	ATTCAGACGA	TTTGGCAGCG	TCCAGGGTAA	GCTGGTGTTA	TAATACGCTG	CTGATCTACA	2820
	TCACAGATTT	GCTAATAAAT	TTACAGTGGG	CCCTGGCATA	TCTCTGTTCA	GTTAGAGTGA	2880
55	GTGCTGACCC	AACAGCCTCT	GTGGTCAAGC	GAGTCACGAA	TGATTATATCA	TAAAGAAAAA	2940
	TCAGTTTTTG	ACTGACCTGG	ATATCCATGA	GCTGCACCTGA	TCACCATGTA	AGGTCACATT	3000
	TAGTAAATGC	TGAAATAAAA	TGATTAAATG	ATTTATCAAT	AAAAGCCTTT	GAAAAACTCT	3060
	TGGATAATAA	ATTTGGAGTTT	TAAAAATGCA	AATTTGCTTA	GTATCTAATA	ATGAAGTGT	3120
	ATTACATATA	GCCGGAATTG	AGGATCTCTT	TGATCCTGGA	AATGGTTTAC	CTAAAAGCTA	3180
60	CAGAACCAGG	CCAATAATAT	TTGAAATATT	GATGCAGACA	AATGAAATAA	TAAAGAGATT	3240
	TTTATGGTTT	ATAAAAATCT	TTTTTGATAT	GATAATAATC	ATGATCACAA	CTGAGATCAA	3300
	AAAAATATAT	GACAGATTAT	TTTGTFTAAA	AATGCAGTTT	TAATTATCTT	AGTCTATAGA	3360
	AATGATCAT	GCATGGAGGC	ATGTATAGGT	ATGATCTGTG	TAAATCTGTA	CATAAAAACA	3420
	GTGCTATTCT	GAGTGAAAAA	TTTTTTGATG	TGCTTACATA	ACCATGGTGA	TTAAAAATGAG	3480
65	TTTATATTTT	TTCTCAAAAA	TTTTAGCAGT	GTGTAAGTA	AGTAATCTTT	AACTGAATCT	3540
	TGACCACTTA	AAAAAAATC	TAAAAATTGA	ACTACCTATA	GTAGTCTGTG	TTTAAAGTGA	3600
	ATTTTAAAG	ACAAAGCAT	CTAAATGAAC	TCAATATAAA	AACATTCATT	TGGAATGTAC	3660
	ATACTGAAAA	ATACAGGTTT	TTTTGACCAA	AAGTTTTTAT	ATCTTTTCTT	TTTATTTATT	3720
	TTTTTCTTAA	GTGCCAACAA	TTTTCTAGAT	ATTATATACA	ACACAGGCTT	TGATCTTGGG	3780
70	GACTTTTCCC	ATATATTTCA	CACTGGAGTG	AATGAAGTTG	TACTTCATTT	CTAGAGAAAA	3840
	GTATATCCCA	GGTCCCCAAT	TGAGAATGTC	TTGCTTGATT	GAAAACGACA	TCATCCCTTG	3900
	GTATACCTCA	GGGATTTGGT	TCAGGACCCC	TGCATTACCC	AAAAATTTGT	CACACTCAAG	3960
	TCCTGCAGTC	ACCCCTGCCT	AAAGATAGAA	TGGCTTCTCT	GTTTTCTTCT	TGAAATACAA	4020
	CCAGAAACAA	TGTGTCTATT	TCTGAAGAA	TAGGATTAAT	GATCATACAA	ATGGGTAAAT	4080
75	CCTGAATTCT	GGTTGTAAAT	CTGGTTACAG	CATAACTAGG	ATTATAATGC	TGCCCTCAATT	4140
	TCACAGCACT	ACTTGCTTAT	ATTGACAACA	AATCATCTCG	CTAAAGAGTG	AATGTAGGCC	4200
	AGGCGGGTGG	GCTCATGCTT	GTAATCCCAG	CACCTTTGGGA	GGCCGAGGCG	GGTGGATCAC	4260
	GAGGTGAGGA	GATCGAGACC	ATCCTGGCTA	ACATGGTAAA	ACCCCGTCTC	TACTAAAAAT	4320
	AGAAAAAAG	AAATTAGCCT	AGCGTGGTGG	CTGGCGGGCG	CCTGTAGTCC	CAGCTATTTG	4380
80	GGAGGCTAAG	CAGGAGGAAT	GGCGTGAACC	CGGGAGGCGG	AGCTTGCAGT	GAGCCGAGGT	4440
	CGTGCCACTG	CACTCCAGCC	TGGGCGACAG	AGCAAGACTC	CGTCTCAAAA	AAAAAATAAA	4500

5
10
15
20
25
30

AAAAAAAAA	AGAGTGAATG	TAATAGTCTT	GCAGAAAATG	AATGAATACC	TTTGTTCAT	4560
AAAGGAAATA	TGCACCTGCT	ACTTTTTTGA	AGGAAATGCC	AAAGTTACGT	TTTACAACAA	4620
GGCTAGAGTT	TGTAAATCTT	GGGTTCAATT	GTGATGACAT	AAGTCAGCAA	ACTGCGGGAA	4680
TACTGTCTCT	TCTATGTATT	TTGTGAATAG	TAAGCATAAT	TTTAGTTTGT	TATTATCAAT	4740
GAAATTTTCA	CTTGAAATTA	AAGCTGCCTT	TTGTTATATT	TTTAACCTAT	AGGATAAGAT	4800
TCCAGTATGT	TATATGAGTT	TTAACAAT	AAAAAATCAA	ATCATGTACA	TTTGAAAAATA	4860
TTTGACACACA	TTTAAAAATA	AATGTAAAGT	TGTCTTTTAA	ACTACTCGGA	TGTGTCCTTT	4920
CTGAACAAAA						

SEQ ID NO:232 PFD4 Protein sequence:

Protein Accession #: O43511

1 11 21 31 41 51

MAAPGGRSEP	PQLPEYSCSY	MVSRPVYSEL	AFQQQHERRL	QERKTLRESL	AKCCSCSRKR	60
AFGVLTLP	ILEWLKPYRV	KEWLLSDVIS	GVSTGLVATL	QGMAYALLAA	VPVGYGLYSA	120
FFPILTYFIF	GTSRHISVGP	FPVVSMLVGS	VVLSMAPDEH	FLVSSSNGTV	LNTTMDTAA	180
RDTARVLAS	ALTLLVGIQ	LIFGGLQIGF	IVRYLADPLV	GGFTTAAAFQ	VLVSQKLKIVL	240
NVSTKNYNGV	LSIIYTLVEI	FQNIQDNLNA	DFTAGLLTIV	VCMAYKELND	RFRHKIPVPI	300
PIEVIVTIIA	TAISYGANLE	KVYNAGIVKS	IPRGFLPPEL	PPVSLFSEML	AASFSAIVVA	360
YAIASVGVKV	YATKYDYTID	GNQEFIAFGI	SNIFSGFFSC	FVATTALSRT	AVQESTGGKT	420
QVAGIISAAI	VMIAILALGK	LLEPLQKSVL	AAVVIANLKG	MFMLQCDIPR	LWRQNKIDAV	480
IWTFCTIVSI	ILGLDLGLLA	GLIFGLLTIV	LRVQFPSPWNG	LGSIPSTDIV	KSTKNYKNIE	540
EPQGVKILRF	SSPIFYGNVD	GFKKCIKSTV	GFDAIRVYNK	RLKALRKIQK	LIKSGQLRAT	600
KNGIISDAVS	TNNAFEPDED	IEDLEELDIP	TKEIEIQVDW	NSELFPKVNV	PKVPIHSLVL	660
DCGAISFLDV	VGVRSLRVIV	KEFQRIDVNV	YFASLQDYVI	EKLEQCGFFD	DNIRKDTFFL	720
TVHDAILYLQ	NQVKSQEGQG	SILETTITLIQ	DKDPTLELIE	TELTEEELDV	QDEAMRTLAS	780
QDEAMRTLAS						

SEQ ID NO:233 PFH2 DNA SEQUENCE:Nucleic Acid Accession #: NM_016029
Coding sequence: 228-1097 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51

CTGCGATCCC	GCAGGGCAGC	GACGCGACTC	TGGTGCGGGC	CGTCTTCTTC	CCCCCGAGCT	60
GGGCGTGCGC	GGCCGCAATG	AACTGGGAGC	TGCTGCTGTG	GCTGCTGGTG	CTGTGCGCGC	120
TGCTCCTGCT	CTTGCTGCAG	CTGCTGCGCT	TCCTGAGGGC	TGACGGCGAC	CTGACGCTAC	180
TATGGGCGGA	GTGGCAGGCA	CGACGCCAG	AAATGGGAGCT	GACTGATATG	GTGGTGTGGG	240
TGACTGGAGC	CTCAGGTGGA	ATTGGTGAGG	AGCTGGCTTA	CCAGTTGTCT	AAACTAGGAG	300
TTTCTCTTGT	GCTGTCAGCC	AGAAGAGTGC	ATGAGCTGGA	AAGGGTGAAG	AGAAGATGCC	360
TAGAGAAATGG	CAATTTTAAA	GAAAAAGATA	TACTTGTTTT	GCCCCCTGAC	CTGACCGACA	420
CTGGTTCCTCA	TGAAGCGGCT	ACCAAAGCTG	TTCTCCAGGA	GTTTGGTAGA	ATCGACATTC	480
TGGTCAACAA	TGGTGAATG	TCCAGCGT	CTCTGTGCAT	GGATACCAGC	TTGGATGTCT	540
ACAGAAAGCT	AATAGAGCTT	AACTACTTAG	GGACGGTGTC	CTTGACAAAA	TGTGTTCTGC	600
CTCACATGAT	CGAGAGGAAG	CAAGGAAAGA	TTGTTACTGT	GAATAGCATC	CTGGGTATCA	660
TATCTGTACC	TCTTTCATT	GGATACTGTG	CTAGCAAGCA	TGCTCTCCGG	GGTTTTTTTA	720
ATGGCCTTCG	AACAGAACTT	GCCACATACC	CAGGTATAAT	AGTTTCTAAC	ATTTGCCCG	780
GACCTGTGCA	ATCAAAATAT	GTGGAGAATT	CCCTAGCTGG	AGAAGTCACA	AAGACTATAG	840
GCAATAATGG	AGACCACTCC	CACAAGATGA	CAACCACTCG	TTGTGTGCGG	CTGATGTTAA	900
TCAGCATGGC	CAATGATTTG	AAAGAAGTTT	GGATCTCAGA	ACAACCTTTC	TTGTTAGTAA	960
CATATTGTGT	GCAATACATG	CCAACCTGGG	CCTGGTGGAT	AACCAACAAG	ATGGGGGAAGA	1020
AAAGGATTGA	GAACCTTAA	AGTGGTGTGG	ATGCAGACTC	TTCTTATTTT	AAAATCTTTA	1080
AGACAAAACA	TGACTGAAAA	GAGCACCTGT	ACTTTTCAAG	CCACTGGAGG	GAGAAATGGA	1140
AAACATGAAA	ACAGCAATCT	TCTTATGCTT	CTGAATAATC	AAAGACTAAT	TTGTGATTTT	1200
ACTTTTAAAT	AGATATGACT	TTGCTTCCAA	CATGGAATGA	AATAAAAAAT	AAATAATAAA	1260
AGATTGCCAT	GAATCTTGCA	AA				

SEQ ID NO:234 PFH2 Protein sequence:

Protein Accession #: NP_057113

1 11 21 31 41 51

MNWEELLWLL	VLCALLLLLV	QLLRFLRADG	DLTLLEAEWQ	GRRPEWELTD	MVVWVTGASS	60
GIGEELEYQL	SKLGVSLVLS	ARRVHELERV	KRRLENGNL	KEKDILVPL	DLTDTGSHEA	120
ATKAVLQEFQ	RIDLIVNNGG	MSQRSLCMDT	SLDVYRKLIE	LNLYLGTVSLT	KCVLPHMIER	180
KQGIKIVTNS	ILGIISVPLS	IGYCASKHAL	RGFFNGLRTE	LATYPGLIIVS	NICPGPVQSN	240
IVENSLAGEV	TKTIGNNGDQ	SHKMTTSRCV	RLMLISMAND	LKEVWISEQP	FLLVTYLWQY	300
MPTWAWITN	KMGKKRIENF	KSGVDADSSY	FKIKFTKHD			

SEQ ID NO:235 ACC5 DNA SEQUENCE

Nucleic Acid Accession #: NM_000450

Coding sequence: 1-1833 (underlined sequences correspond to start and stop codons)

```

1      11      21      31      41      51
5      |      |      |      |      |      |
      ATGATTGCTT CACAGTTTCT CTCAGCTCTC ACTTTGGTGC TTCTCATTAA AGAGAGTGGG 60
      GCCTGGTCTT ACAACACCTC CACGGAAGCT ATGACTTATG ATGAGGCCAG TGCTTATTGT 120
      CAGCAAAGGT ACACACACCT GGTTCGAATT CAAAACAAAG AAGAGATTGA GTACCTAAAC 180
      TCCATATTGA GCTATTACCC AAGTTATTAC TGGATTGGAA TCAGAAAAGT CAACAATTGT 240
      TGGGTCTGGG TAGGAACCCA GAAACCTCTG ACAGAAGAAG CCAAGAACTG GGCTCCAGGT 300
10     GAACCCAACA ATAGGCAAAA AGATAGGAGC TGCCTGGAGA TCTACATCAA GAGAGAAAAA 360
      GATGTGGGCA TGTGGAATGA TGAGAGGTGC AGCAAGAAGA AGCTTGCCCT ATGCTACACA 420
      GCTGCCTGTA CCAATACATC CTGCAGTGGC CACGGTGAAT GTGTAGAGAC CATCAATAAT 480
      TACACTTGCA AGTGTGACCC TGGCTTCAGT GGACTCAAGT GTGAGCAAAT TGTGAACTGT 540
      ACAGCCCTGG AATCCCTGTA GCATGGAAGC CTGGTTTGCA GTCACCCACT GGGAACTTC 600
15     AGCTACAATT CTTCCTGCTC TATCAGCTGT GATAGGGGTT ACCTGCCAAG CAGCATGGAG 660
      ACCATGCAGT GTATGTCTCT TGGAGAATGG AGTGCTCCTA TTCCAGCCTG CAATGTGGTT 720
      GAGTGTGATG CTGTGACAAA TCCAGCCAAT GGGTTCGTGG AATGTTTCCA AAACCTTGA 780
      AGCTTCCCAT GGAACACAAC CTGTACATTT GACTGTGAAG AAGGATTTGA ACTAATGGGA 840
      GCCAGAGGCC TTCAGTGTAC CTCATCTGGG AATTGGGACA ACGAGAAGCC AACGTGTAAA 900
20     GCTGTGACAT SCAGGGCCGT CCGCCAGCCT CAGAAATGGCT CTGTGAGGTG CAGCCATTCC 960
      CCTGCTGGAG AGTTCACTTT CAAATCATCC TGCAACTTCA CCTGTGAGGA AGGCTTCATG 1020
      TTGCAGGGAC CAGCCAGGT TGAATGCACC ACTCAAGGCG AGTGGACACA GCAAATCCCA 1080
      GTTTGTGAAG CTTTCCAGTG CACAGCCTTG TCCAACCCCG AGCGAGGCTA CATGAATTGT 1140
      CTTCCTAGTG CTTCCTGCGC TTTCCTTTAT GGGTCCAGCT GTGAGTTCTC CTGTGAGCAG 1200
25     GGTTTTGTGT TGAAGGGATC CAAAAGGCTC CAATGTGGCC CCACAGGGGA GTGGGACAA 1260
      GAGAAGCCCA CATGTGAAGC TGTGAGATGC GATGCTGTCC ACCAGCCCCC GAAGGGTTTG 1320
      GTGAGGTGTG CTCATTCCCC TATTGGAGAA TTCACCTACA AGTCTCTCTG TGCCTTCAGC 1380
      TGTGAGGAGG GATTTGAAAT ATATGGATCA ACTCAACTTG AGTGACATC TCAGGGACAA 1440
      TGGACAGAAG AGGTTCTCTC CTGCCAAGTG GTAAAATGTT CAAGCCTGGC AGTTCGGGGA 1500
30     AAGATCAACA TGAGCTGCAG TGGGGAGCCC GTGTTTGGCA CTGTGTGCAA GTTCGCTGT 1560
      CCTGAAGGAT GGACGCTCAA TGGCTCTGCA GCTCGGACAT GTGAGGCCAC AGGACACTGG 1620
      TCTGGCCTGC TACCTACCTG TGAAGCTCCC ACTGAGTCCA ACATTCCTCT GGTAGCTGGA 1680
      CTTCCTGCTG CTGGACTCTC CCTCCTGACA TTAGCACCAT TTCTCCTCTG GCTTCGGAAA 1740
      TGCTTACGGA AAGCAAAGAA ATTTGTTCTT GCCAGCAGCT GCCAAAGCCT TGAATCAGAC 1800
35     GGAAGCTACC AAAAGCCTTC TTACATCCTT TAA

```

SEQ ID NO:236 ACC5 Protein sequence:

Protein Accession #: NP_000441

```

40
1      11      21      31      41      51
45     |      |      |      |      |      |
      MIASQFLSAL TLVLLIKESG AWSYNTSTEA MTYDEASAYC QQRVTHLVAI QNKEEIEYLN 60
      SILSYSPSYW WVGIRKVNIV WVVGTQKPL TEEAKNWAPG EPNNRQKDED CVEIYIKREK 120
      DVGMMWDERC SKKKLALCYT AACTNTSCSG HGECVETINN YTCKCDPFSS GLKCEQIVNC 180
      TALESPHGS LVCSEPLGNF SYNSSCSISC DRGYLPSSME TMQCMSSGEW SAPIPACNVV 240
50     ECDAVTNPAN GEVECFQNPQ SFPWNTTCTF DCEEGFELMG AQLQCTSSG NWDNEKPTCK 300
      AVTCRAVRQP QNGSVRCSSH PAGEFTFKSS CNFTCEEGFM LQGPQVETL TQGWTTQIIP 360
      VCEAFQCTAL SNPERGYMNC LPSASGSFRY GSSCEFSCEQ GFVLKGSKRL QCGPTGEWDN 420
      EKPTCEAVRC DAVHQPPKGL VRCASPIGE FTYKSSCAFS CEEGFELYGS TQLECTSQGQ 480
      WTEEVFSCQV VKCSSLAVPG KINMSCSGEP VFGTVCKFAC PEGWTLNGSA ARTCGATGHW 540
55     SGLLPTCEAP TESNIPLVAG LSAAGLSLLT LAPFLWLRLK CLRKAKKFPV ASSCQSLESD 600
      GSYQKPSYLL

```

SEQ ID NO:237 PM28 DNA SEQUENCE

Nucleic Acid Accession #:

N51002

Coding sequence: 1-3793 (underlined sequences correspond to start and stop codons)

```

60
1      11      21      31      41      51
      ATGATGTGTG AAGTGATGCC CACGATTAAT GAGGACACCC CAATGAGCCA AAGGGGGTCC 60
      CAAAGCAGTG GCTCGGACTC AGACTCCCAT TTTGAGCAGC TGATGGTGAA TATGCTAGAT 120
65     GAAAGGGATC GTCTTCTAGA CACCTTCGG GAGACCCAGG AAAGCCTCTC ACTTGCCAG 180
      CAAAGACTTC AGGATGTCTC CTATGACCGA GACTCACTCC AGAGACAGCT CAATTTCAGC 240
      CTGCCACAGG ATATCGAATC CCTAACAGGA GGGCTGGCTG GTTCTAAGGG GGCTGATCCA 300
      CCGGAATTTG CTGCACTGAC AAAAGAATTA AATGCCTGCA GGGAACAACT TCTAGAAAAG 360
      GAAGAAGAAA TCTCTGAAC TAAAGCTGAA AGAAACAACA CAAGACTATT ACTGGAGCAT 420
70     TTGGAGTGCC TTGTGTACAG ACATGAAAGA TCACTAAGAA TGACGGTGGT AAAACGGCAA 480
      GCCAGTCTC CCTCAGGAGT ATCCAGTGAA GTTGAAGTTC TCAAGGCACT GAAATCTTTG 540
      TTTGAGCACC ACAAGGCCCTT GGATGAAAAG GTAAGGGAGC GACTGAGGGT TTCTTTAGAA 600
      AGAGTCTCTG CACTGGAAGA AGAAGTAGCT GCTGCTAATC AGGAGATTGT TGCTTGCCT 660
      GAACAAAATG TTCATATACA AAGAAAATG GCATCAAGCG AGGGATCCAC AGAGTCAGAA 720
75     CATCTTGAAG GGATGGAACC TGGACAGAAA GTCCATGAGA AGCGTTTGTG CAATGGTTCT 780
      ATAGACTCAA CCGATGAAAC TAGTCAAATA GTTGAACATC AAGAATTGCT TGAAGAAGCAA 840
      AACTATGAAA TGGCCAGATC GAAAGAAGCT TTAGCAGCCC TTTCTTCCCG AGTGGGAGAG 900
      GTGGAACAGG AAGCAGAGAC AGCAAGAAAG GATCTCATTA AAACAGAAGA AATGAACACC 960
      AAGTATCAA GGGACATTAG GGAGGCCATG GCACAAAAGG AAGATATGGA AGAAGAATTT 1020
80     ACAACCCTTG AAAAGCGTTA CCTCAGTGCT CAGAGAGAAT CTACCTCCAT ACATGACATG 1080

```

	AATGATAAAC	TAGAAAAATGA	GTTAGCAAAT	AAAGAAGCTA	TCCTACGGCA	GATGGAAGAG	1140
	AAAAACAGAC	AGTTACAAGA	ACGTCTTGAG	CTAGCTGAAC	AAAAGTTGCA	GCAGACCATG	1200
	AGAAAGGCTG	AAACCTTGCC	TGAAGTAGAG	GCTGAACTGG	CTCAGAGAAT	TGCAGCCCTA	1260
5	ACCAAGGCTG	AAGAGAGACA	TGGAAATATT	GAAGAACGTA	TGAGACATTT	AGAGGGTCAA	1320
	CTTGAAGAGA	AGAATCAAGA	ACTTCAAAGA	GCTAGGCAAA	GAGAGAAAAT	GAATGAGGAG	1380
	CATAACAAGA	GATTATCGGA	TACGGTTGAT	AGACTTCTGA	CTGAATCCAA	TGAACGCCTA	1440
	CAACTACACT	TAAAGGAAAG	AATGGCTGCT	CTAGAAGAAA	AGAATGTTTT	AATTCAAGAA	1500
	TCAGAAACTT	TCAGAAAGAA	TCCTGAAGAA	TCTTTACATG	ATAAGGAAAAG	ATTAGCAGAA	1560
10	GAAATTGAAA	AGCTGAGATC	TGAACCTGAC	CAATTGAAAA	TGAGAACTGG	CTCTTTAATT	1620
	GAACCCACAA	TACCAAGAAC	TCATCTAGAC	ACCTCAGCTG	AGTTGCGGTA	CTCAGTGGGA	1680
	TCCTTAGTGG	ACAGCCAGTC	TGATTACAGA	ACAACATAAG	TAATAAGAAG	ACCAAGGAGA	1740
	GGCCGCATGG	GTGTGCGAAG	AGATGAGCCA	AAGGTGAAAT	CTCTTGGGGA	TCACGAGTGG	1800
	AATAGAACTC	AACAGATTGG	AGTACTAAGC	AGCCACCCTT	TTGAAAGTGA	CACTGAAATG	1860
15	TCGTGATTTG	ATGATGATGA	CAGAGAAACA	ATTTTTAGCT	CAATGGATCT	TCTCTCTCCA	1920
	AGTGGTCATT	CCGATGCCCA	GACGCTAGCC	ATGATGCTTC	AGGAACAATT	GGATGCCATC	1980
	AACAAAGAAA	TCAGGCTAAT	TCAGGAAGAA	AAAGAATCTA	CAGAGTTGCG	TGCTGAAGAA	2040
	ATTGAAAAATA	GAGTGGCTAG	TGTGAGCCTC	GAAGGCCTGA	ATTTGGCAAG	GGTCCACCCA	2100
	GGTACCTCCA	TFACTGCCCT	TGTTACAGCT	TCATCGCTGG	CCAGTTCAATC	TCCCCCAGT	2160
20	GGACACTCAA	CTCCAAAGCT	CACCCCTCGA	AGCCCTGCCA	GGGAAATGGA	TCCGATGGGA	2220
	FTCATGACAC	TGCCAAGTGA	TCTGAGGAAA	CATCGGAGAA	AGATTGCACT	TGTGGAAGAA	2280
	GATGGTGCAG	AGGACAAAGC	AACAATTAAA	TGTGAAACTT	CTCCTCCTCC	TACCCCTAGA	2340
	GCCTCAGAA	TGACTCACAC	TCCTCCCTCT	TCCTACCACA	ATGATGCTCG	AAGTAGTTTA	2400
	TCTGTCTCTC	TTGAGCCAGA	AAGCCTCGGG	CTTGGTAGTG	CCAACAGCAG	CCAAGACTCT	2460
	CTTCACAAAG	CCCCCAAGAA	GAAAGGAATC	AAGTCTTCAA	TAGGACGTTT	GTTTGGTAAA	2520
25	AAAGAAAAAG	CTCGACTTGG	GCAGCTCCGA	GGCTTTATGG	AGACTGAAAGC	TGCAGCTCAG	2580
	GAGTCCCTGG	GGTTAGGCAG	ACTCGGAACT	CAAGCTGAGA	AGGATCGAAG	ACTAAAGAAA	2640
	AAGCATGAAC	TTCTTGAAGA	AGCTCGGAGA	AAGGGATTAC	CTTTTGCCCA	GTGGGATGGG	2700
	CCAACGTGTT	TCGCATGGCT	AGAGCTTTGG	TTGGGAATGC	CTGCGTGGTA	CGTGGCAGCC	2760
30	TGCCGAGCCA	ACGTGAAGAG	TGGTGCCATC	ATGCTGCTCT	TATCTGACAC	TGAGATCCAG	2820
	AGAGAAATFG	GAATCAGCAA	TCCACTGCAAT	CGCTTAAATC	TTTGATTAGC	AATCCAGGAG	2880
	ATGGTTTCCC	TAACAAGTCC	TTTCACTCCT	CCAACATCTC	GAACCTCTTC	AGGCAACGTT	2940
	TGGGTGACTC	ATGAAGAAAT	GGAAAATCTT	GCAGCTCCAG	CAAAAACGAA	AGAATCTGAG	3000
	GAAGGAAGCT	GGGCCAGTGG	TCCGGTTTTT	CTACAGACCC	TGGCTTATGG	AGATATGAAT	3060
35	CATGAGTGGG	TTGGAATAGA	ATGGCTTCCC	AGCTTGGGGT	TACCTCAGTA	CAGAAGTTAC	3120
	TTTATGGAAT	GCTTGGTAGA	TGCAAGAAATG	TTAGATCACC	TAACAAAAAA	AGATCTCCGT	3180
	GTCCATTTAA	AAATGGTGGG	TAGTTTCCAT	CGAACAAGTT	TACAATATGG	AATTATGTGC	3240
	TTAAAGAGGT	TGAATTTATG	CAGAAAGAAA	CTAGAAAGAA	GACGGGAAGC	AAGCCAACAT	3300
	GAAATATAAG	ACGTGTTGGT	TGGGAGCAAT	GACCGAATTA	TTTCCTGGAT	ACAAGCAATT	3360
40	GGACTTCGAG	AATATGCAAA	TAAATATACT	GAGAGCGGTG	TGCATGGCTC	ACTTATAGCC	3420
	CTGGATGAAA	ACTTTGACTA	CAGCAGCTTA	ACTTTATTAT	TACAGATTCC	AACACAGAAC	3480
	ACCCAGGCAA	GGCAGATCTT	TGAAGAAGAA	TACAATAACC	TCTTGGCCCT	GGGAAGTGAA	3540
	AGGGGACTGG	ATGAAAGTGA	TGACAAGAAC	TTTCAAGCTG	GATCAACCTG	GAGAAGGCAG	3600
	TTTCTCTCTC	GTGAAGTACA	TGGAATCAGC	ATGATGCCTG	GGTCTCTAGA	AACATTACCA	3660
45	GCTGGATTTA	GGTTAACCAC	AACCTCTGGG	CAATCAAGAA	AAATGACAAC	AGATGTTGCT	3720
	TCATCAAGAC	TGCAGAGGTT	AGACAACCTC	ACTGTTTCGA	CATACTCATG	TCTCGAGTAA	3780
	GCGGCCGCTT	TAA					

SEQ ID NO:238 PM28 Protein sequence:

Protein Accession #: none found

	1	11	21	31	41	51	
55	MMCEVMTPTN	EDTPMSQRGS	QSSGSDSDSH	FEQLMVNMLD	ERDRLLDTLR	ETQESLSLAQ	60
	QRLQDVIYDR	DSLQRLNSA	LPQDIESLFG	GLAGSKGADP	PEFAALTKEL	NACREQLLEK	120
	EEIESELKAE	RNNTRLLLEH	LECLVSRHER	SLRMTVVVRQ	AQSPSGVSSE	VEVLKALKSL	180
	FEHHKALDEK	VRELRVLSLE	RVSALIEEELA	AANQEIVLR	EQNVIHQKRM	ASSEGSTESE	240
	HLEGMEPGQK	VHEKRLSNGS	IDSTDETSQI	VELQELLEKQ	NYEMAQMKER	LAALSSRVGE	300
60	VEQEAETARK	DLIKTEEMNT	KYQDIREAM	AQKEDMEERI	TTLEKRYLSA	QRESTSIHDM	360
	NDKLENELAN	KEAILRQMBE	KNRQLQERLE	LAEQKLQQTM	RKAETLPEVE	AELAQRIAAL	420
	TKAERHNGNI	EERMRHLEQG	LEEKHQELQR	ARQREKMNEE	HNKRLSDTVD	RLLTESNERL	480
	QLHLKERMAA	LEEKNVLIQE	SETFRKNLEE	SLHDKERLAE	EIEKLRLSEL	QLKMRGTSLI	540
	EPTIPRTHLD	TSALRYSVG	SLVDSQSDYR	TTKVIRRPFR	GRMGVRRDEP	KVKSLGDHEW	600
65	NRTQQIGVLS	SHPFESDTEM	SDIDDDRET	IFSSMDLLSP	SGHSDAQTLA	MMLQEQLDAI	660
	NKEIRLIQEE	KESTELRAEE	IENRVASVSL	EGLNLARVHP	GTSITASVTA	SSLASSPPPS	720
	GHSTPKLTFR	SPAREMDRMG	VMTLPDLRKL	HRRKIAVVEE	DGREDKATIK	CETSPPTPR	780
	ALRMTHTLPS	SYHNDARSSL	SVSLEPESLG	LGSANSSQDS	LHKAPKKKGI	KSSIGRLFGK	840
	KEKARLGQLR	GFMEFEAAAQ	ESLGLGKLGT	QAEKDRRLKK	KHELLEEARR	KGLPFAQWDG	900
70	PTVVAVLELW	LGMPAWYVAA	CRANVKSAGI	MSALSDTEIQ	REIGISNPLH	RLKLRLAIQE	960
	MVSLTSPSAP	PTSRTPSGNV	VWTHEEMENL	AAPAKTKESE	EGSWAQCFVF	LQTLAYGDMN	1020
	HEWIGNEWLP	SLGLPQYRSY	FMECLVDARM	LDHLTKKDLR	VHLKMDVSFH	RTSLQYGIMC	1080
	LKRLNYDRKE	LERRREASQH	EIKDVLVWSN	DRIIRWIQAI	GLREYANNIL	ESGVHGSLLA	1140
75	LDENFDYSSL	TLLQLIPTQN	TQARQILERE	YNNLLALGTE	RRLESDDDKN	FRRGSTWRRQ	1200
	FPFREVHGIS	MMPGSSETLP	AGFRLTTTSG	QSRKMTTDVA	SSRLQRLDNS	TVRTYSCLLE	

SEQ ID NO:239 PCI4 DNA SEQUENCE

Nucleic Acid Accession #: NM_016570

Coding sequence: 1- 1134 (underlined sequences correspond to start and stop codons)

80

1 11 21 31 41 51
 5 ATGAGGCGAC TGAATCGGAA AAAAAGTTTA AGTTTGGTAA AAGAGTTGGA TGCCTTTCCG 60
 AAGGTTCCCTG AGAGCTATGT AGAGACTTCA GCCAGTGGAG GTACAGTTTC TCTAAATAGCA 120
 TTTACAACATA TGGCTTTATT AACCATAATG GAATTCCTCAG TATATCAAGA TACATGGATG 180
 AAGTATGAAT ACGAAGTAGA CAAGGATTTT TCTAGCAAAAT TAAGAATTAA TATAGATATT 240
 ACTGTTGCCA TGAAGTGTCA ATATGTTGGA GCGGATGTAT TGGATTTAGC AGAAACAATG 300
 10 GTTGCATCTG CAGATGGTGT AGTTTATGAA CCAACAGTAT TTGATCTTTC ACCACAGCAG 360
 AAAGAGTGGC AGAGGATGCT GCAGCTGATT CAGAGTAGGC TACAAGAAGA GCATTCACTT 420
 CAAGATGTGA TATTTAAAG TGCTTTTAAA AGTACATCAA CAGCTCTTCC ACCAAGAGAA 480
 GATGATTGAT CACAGTCTCC AAATGTCATG AGAATTCATG GCCATCTATA TGTCATAAAA 540
 GTAGCAGGGA ATTTTCACAT AACAGTGGGC AAGGCAATTC CACATCCTCG TGGTCATGCA 600
 CATTTGGCAG CACTTGTCAA CCATGAATCT TACAATTTT CTATAGAAAT AGATCATTTG 660
 15 TCTTTTGGAG AGCTTGTTC AGCAATTATT AATCCTTTAG ATGGAAGTGA AAAAATTGCT 720
 ATAGATCACA ACCAGATGTT CCAATATTTT ATTACAGTTG TGCCAACAAA ACTACATACA 780
 TATAAAATAT CAGCAGACAC CCATCAGTTT TCTGTGACAG AAAGGGAACG TATCATTAAC 840
 CATGCTGACG GCAGCCATGG AGTCTCTGGG ATATTTATGA AATATGATCT CAGTTCTCTT 900
 20 ATGTTGACAG TTACTGAGGA GCACATGCCA TTCTGGCAGT TTTTGTGAAG ACTCTGTGGT 960
 ATTTGTTGGAG GAATCTTTTC AACAACAGGC ATGTTACATG GAATTGGAAT ATTTATAGTT 1020
 GAAATAATTT GCTGTCTTT CAGACTTGA TCCTATAAAC CTGTCAATTC TGTTCCTTTT 1080
 GAGGATGGCC ACACAGACAA CCACTTACCT CTTTTAGAAA ATAATACACA TTGA

25 SEQ ID NO:240 PC14 Protein sequence:
 Protein Accession #: NP_057654

1 11 21 31 41 51
 30 MRLNRKKTL SLVKELDAFP KVPESYVETS ASGGTVSLIA FTTMALLTIM EFSVYQDTWM 60
 KYEYEVDDKDF SSKLRINIDI TVAMKCOYVG ADVLDLAETM VASADGLVYE PTVFDLSPQQ 120
 KEWQRLQLI QSRLEEHSL QDVIFKSAFK STSTALPRE DDSSQSPNAC RIHGHLYVKN 180
 VAGNFHITVG KAIHPRGHA HLAALVNHE S YNFSHRIDHL SFGELVPAII NPLDGTKEIA 240
 35 IDHNMQFYF ITVVPFKLHT YKISADTHQF SVTERERIIIN HAAGSHGVSG IFMKYDLSSL 300
 MVTVTEEHMP FWQFFVRLCG IVGGIFSTTG MLHGIGKFIV EIICCRFRLG SYKPVNSVFP 360
 EDGHTDNLHP LLENNTH

SEQ ID NO:241 PBA7 DNA SEQUENCE

40 Nucleic Acid Accession#: AA219134
 Coding sequence: 24-1815 (underlined sequences correspond to start and stop codons)

AATTCGCCCT TGCTTAATTA AGCATGTTTA CCTTCCTGTC ATCTGTCCT GCTGCTGTCA 60
 GTGGCCTCCT GGTGGGTTAT GAAGTTGGGA TCATCTCTGG GGCTCTTCTT CAGATCAAAA 120
 45 CCTTATTAGC CTTGAGCTGC CATGAGCAGG AAATGGTTGT GAGCTCCCTC GTCATTGGAG 180
 CCTCCTTGG CTCACCTACC GGAGGGGTC TGATAGACAG ATATGGAAGA AGGACAGCAA 240
 TCATCTGTG ATCCTGCTG CTGGAGCTCG GAAGCTTAGT CTGTATCCTC AGTTTATCCT 300
 ACACGGTTCT TATAGTGGGA CGATTGCCA TAGGGGTTTC CATCTCCCTC TCTTCCATTG 360
 50 CCACTGTGTT TTACATCGCA GAGATTGCTC CTCAACACAG AAGAGGCCTT CTGTGTCTAC 420
 TGAATGAGCT GATGATTGTC ATCGGCATTC TTCTGCCTA TATTTCAAAAT TACGCATTG 480
 CCAATGTTTT CCATGGCTGG AAGTACATGT TTGGTCTTGT GATTCCTTGG GGAGTTTGGC 540
 AAGCAATTGC AATGATTTT CTTCCTCAA GCCCTCGGT TCTGGTGATG AAAGGACAAG 600
 AGGGAGCTGC TAGCAAGGTT CTGGAAAGGT TAAGAGCACT CTCAGATACA ACTGAGGAAC 660
 55 TCACGTGTAT CAAATCTCTC CTGAAGATG AATATCAGTA CAGTTTTTGG GATCTGTTTC 720
 GTTCAAAGA CAACATCGCG ACCCGAATAA TGATAGGACT AACACTAGTA TTTTGTGAC 780
 AAATCACTGG CCAACCAAA ATATTGTTCT ATGCATCAAC TGTTTGAAG TCAGTTGGAT 840
 TTCAAAGCAA TGAGGCAGCT AGCCTCGCT CCACTGGGGT TGGAGTCGTC AAGGTCATTA 900
 GCACCATCCC TGCCACTCTT CTGTAGACC ATGTCGGCAG CAAAACATTC CTCTGCATTG 960
 60 GCTCCTCTGT GATGGCAGCT TCGTTGGTGA CCATGGGCAT CGTAAATCTC AACATCCACA 1020
 TGAACATCAC CCATATCTGC AGAAGCCACA ATTCTATCAA CCAGTCCITG GATGAGTCTG 1080
 TGATTATGAG ACCAGGAAAC CTGTCAACCA ACAACAATAC TCTCAGAGAC CACTTCAAAG 1140
 GGATTTCTTC CCATAGCAGA AGCTCACTCA TGCCCTGAG AAATGATGTG GATAAGAGAG 1200
 GGGAGACGAC CTCAGCATCC TTGCTAAATG CTGGATTAAG CCACACTGAA TACCAGATAG 1260
 65 TCACAGACC TGGGAGCTC CCAGCTTTT TGAAATGGCT GTCCTTAGCC AGCTTGCTTG 1320
 TTTATGTTGC TGCTTTTCA ATTGGTCTAG GACCAATGCC CTGGCTGGTG CTCAGCGAGA 1380
 TCTTCTCTGG TGGGATCAGA GGACGAGCCA TGCTTTAAAC TTCTAGCATG AACTGGGGCA 1440
 TCAATCTCCT CATCTCGCTG ACATTTTGA CTGTAAGTGA TCTTATTGGC CTGCCATGGG 1500
 TGTGCTTTAT ATATACAATC ATGAGTCTAG ATCTTATTGG CTGCCATGG GTGTGCTTTA 1560
 70 TATATACAAT CATGAGTCTA GCATCCCTGC TTTTGTGTTG TATGTTTATA CCTGAGACAA 1620
 AGGGATGCTC TTGGGAACAA ATATCAATGG AGCTAGCAAA AGTGAAGTAT GTGAAAAACA 1680
 ACATTTGTTT TATGATGATC CACCAAGAAG AATTAGTGCC AAAACAGCCT CAAAAAGAA 1740
 AACCCAGGA GAGCTCTTGG GAGTGAACA AGCTGTGTGG TAGGGGCCAA TCCAGGCAGC 1800
 75 TTTCTCCAGA GACCTAATGG CCTCAACACC TTCTGAACGT GGATAGTGCC AGAACACTTA 1860
 GGAGGGGTGC TTGGACCAA TGCATAGTTG CGACTCCTGT GCTCTCTTT CAGTGTCTATG 1920
 GAACGTGTTT TGAAGATGAC CTCTGAAATG ATAAAGACAG CTTTAAATCC CCTCCTCMC 1980
 CAGAAGGAAC CTCAAAGGT AGATGAGGTA CAAGGTCTTA AGTGATCTCT TTTCTGAGC 2040
 AGGATATCAG GTTAAAAAAA AAAAGTTACT GGCTGGTTTA ATACTTCTA CCTTCTTCAC 2100
 80 AGAGCAGCCT TTGAATAGAC TATGCTCTAG TGAAGACATC AACCTCCGCC TTAAGCTATG 2160
 TATGATGGA GGCCAGTCG AGCTTATTA TGCAGACACA CAAGTGGTCT GGACATGAGG 2220
 GTACAGTTTC TGCTACCAA GACACTACTT GCACTGGATC TTACGCAAAA AAGAACCAGA 2280

ACACACAGTG TGGACAACTG CCCATATATT CTATCTAGAT TAGGAGAGGG TCCTGGCTAG 2340
 GATTTTAGTG GTAATTCCTA GTTACATTCA ACAAGTATAA AGATTATAGA GCTTATTTTA 2400
 TGAACATAA ACTATAITTT AATGCAAAAT ATCCTTTTAT GAATTTTCATG TTAATATTGT 2460
 GAAATATTAA AATAATTCCT CAATAGTTGA GAAAAATGAG CATTITTTTC CATTITTTAAA 2520
 AAATGCATAG AAAAGACAAT TTAAAAATCC TGGGACCATA TTTATTTAGA AGTAGCTGTT 2580
 AGTAAACAT TAGAAAAGGA GTCAGGCCAT TAGGTTATTT ATCCAAATCT CTAAGCAATT 2640
 AGGTTGAAGT TATTAAGTCA AGCCTAGAAA AGCTGCCTCC TTGTAAGGCT TTCATGACAA 2700
 TGTATAGTAA TCCACAGTGT CCAATTCCTC AACTCCTCA GGAATATCAC TACCTCAGGT 2760
 TACGGTACAC AGGCTATAAT TGATGATGAT GTTCAGATAA CTGAAGACAC AATAAATGAC 2820
 ATTCAGACAT CAGGAMAAWW CCCTCATGTT CTTTCTATG ATGGCCACCT GTACCAGCAA 2880
 CGTGGGTTTC ACCCACACAA CGATGAACTG TTCTCTTACT TCTCCAGTTG ATTTTAAAGA 2940
 CTTGTTAAGA GGCTTTACTA ATAAAAATTG GGTATGATAG AAAAWCCACA ATCAAAWCTT 3000
 GAACCAATA ACATATTAAA TTAATAATAT TTAAGTGATG GAAGACACAC AAAAAACTTA 3060
 AAAGCACGAA CAACCTAAT TGAAAAAGAA TTTTAAATA TGATTAACCT GAAGAAAAGA 3120
 GAATCCTAAG AGCCAAAGCT CCTTTTATT TAGCTTGGAA TTTTCTATT GGTTCCTAAC 3180
 AAATGTCCC AATGTCATAT AAGGAAACAT GATCTATTAC ATTCTTTAT AACAATGTGG 3240
 AGAGACTATA AACCTATGTA AGTAGTAAAA CTATATYAGA GACTCAGGAG ACTGACTAAA 3300
 AGGCCTGGAT CTGCAGTGTG TTATCTGTAT AAAAATTGGC AGGGGGAAGC TAAAGGAAA 3360
 GGAGATTGGA GATCTCAATT CTATCATGGT GTATTTTATA CGAAATCAG AGCATGCATT 3420
 GTTTTGTGTT TTGGAAAAGA GAAGGGAAGT GTGTTCTGCC CCATGTTTCC TTCCGTGTTT 3480
 ATAGTICAAA CTCTATATAT ACTTCAGGTA TTTTGTGTT AGCCCTTCAT TATAAATGGG 3540
 CAGGAAATTG TTTTCAACC TAGCCAGTTT ATTACTAGTG ACCTTGACTT CAGTATCTTG 3600
 AGCATTCITT TATATTTTTC TTTTATTATC CTGAGTCTGT AACTAAACAA TTTTGTCTTC 3660
 AAATTTTAT CCAATATCCA TTGCACCACA CCAATCAAG CTCTTGATT TTCAAAAAATA 3720
 AAAAGGGGGA AATACTTACA ACTTGATACAT ATATATTAC AGTTTATTAT TATAAAAAAA 3780
 ATTTACAGTA CTTATGGAGA GCCAGCAGAA GACATCAGAG CACTCACTTC TTCCCATCTT 3840
 TGTTAAGGTT AGCGAATTAC CCATGGACAC TGTTAGGTGA GGCTCATTCG GCAGCCCTGA 3900
 AAACAAACCT GGTACACACTG TCTTTACCTT CTCCCTTCAG ATAAAGCACT TCGATTATCT 3960
 ATTGATCTGC CCAGTTTICA AGTCATGCGA ATACTAAAAA GGTTACATCA TCTGGATCTG 4020
 TACCTTGGCT ATATAAGCAT GTTTCCCTT TATCTATGT TCTTTTATT GGTGAACATT 4080
 GAAAAACAGG AGGTGACTTA TTACTGTAA TTAATACTAA ATGAAAAATG TCAAGTCTTT 4140
 AAAACAGTGA GCTTGTAACCT CTTTCATGTA ATTTTATTCT CTATGAATTT GGCTATCCTA 4200
 CTGAATCTTA AATAAAAGGA AATAAACACT TTTTITWAA AAAAAAGGAA AAATAMAARW 4260
 MWAAAAATCT CAATGAAATA TTTCACAAGA AGGAAAAA

Protein Accession #: AAF91431

SEQ ID NO:242 PBA7 Protein sequence:

MFTFLSSVTA AVSGLLVGYE LGHSGALLQ IKTLALSCH EQEMVVSSLV IGALLASLTG 60
 GVLIDRYGRR TAILSSCLL GLGSLVLIS LSYTVLIVGR IAIGVSISLS SIATCVYIAE 120
 IAPQHRRGLL VSLNLMIVI GILSAYISNY AFANVFHGWK YMFGLVIFLG VLQAIAMYFL 180
 PPSRFLVMK QGEGAAASKVL GRLRLSDTT EELTVIKSSL KDEYQYSFWD LFRSKDNMRT 240
 RIMIGLTLVF FVQITGQPNL LFYASTVLKS VGFQNEAAS LASTGVGVVK VISTIPATLL 300
 VDHVGSKTFL CIGSSVMAAS LVTMGIIVNLN IHMNFTHICK SHNSINQSLD ESVIYGPGNL 360
 STNNNTLRDH FKGISHSRS SLMPLRNDVD KRGETTSASL LNAGLSHTEY QIVTDPGDVP 420
 AFLKWLSLAS LLVYVAAFSL GLGMPWLVL SEIFPGGIRG RAMALTSSMN WGINLLISLT 480
 FLTVDLIGL PWVCFIYTIM SLDLIGLPWV CFITYIMSLA SLLFVVMFIP ETKGCSLEQI 540
 SMELAKVNYV KNNICFMSHH QEELVPKQPQ KRKPQEQLLE CNKLCGRGQS RQLSPET

SEQ ID NO:243 PAB4 DNA sequence:

Nucleic Acid Accession#: AA172056

Coding sequence: 121-339 (underlined sequences correspond to start and stop codons)

TTTAGCCACC AGAGGANTTC TCTTGAAATA CCAAAATCC ATCAGTATCT TGAATCATGC 60
 TGGATTTTGA AGAATTCCTA AGAAGCCATG TAAAGGGGGC TCTCTGGCCT TGAATAGTG 120
 ATGTTTTTTA TACAGAAAGG AGAATGCAGA ATGGTCAGAC TATCATGCAC TGTAAATTT 180
 GATTICAAGA AATTACAGGA AAACCTTCCA AAGTTCCATC TCACAGAANN TTATTTTNCC 240
 AAGAATTCCTA AGATAAGTTT AGTTTATGG AAGACTTTTA TGTGGTTTTT ACTCACTCTT 300
 CATCTCAGAC ATCGACAGAT GATTACATCA CTTATAGTTC TAGTAAATTT ATTAATATAA 360
 AACTCAGAGA CATTTCAATA TCCACATTGC TTACACCATT AGGCATAGAT TCAGTGTGAG 420
 CTATGACAAT TGAATAAGAG CTGTTTGTG ATTTAAAGGT TTAATTTCT CTAACCAAAC 480
 TGCTTGATCC AGATGCAGGA CTGCAATGT TAATATTGT TCTGGAAGAA CAATCAAATA 540
 AGACTTAAGA GGAAAGGGAA TGGCCACAAT CCACCTGAAA TTTTCTCTTA AAAAGTGTGC 600
 AGCCTACTAA ATCAGAAATGA AATAGAAAGT ACAAGATTAT AAACAAAATG CAATCAAAC 660
 TTTCTTAAGC TTACCTAAAG TTATTTTATC TGAATAATT AAGCAACTTT GTTCAACATT 720
 AAATGACAA TCTAAACTAA CAAGTCTTT GAATTATGC ATGGTAGTAA ACATTCTCTC 780
 TATTAACCTT ATTACCTAAG GCTAAACCTA AAATTTTAA GCAAAATTAG AAAAAATGTC 840
 TCACTCATC AAAAAATAAA GTTGTGTA TTTAGTATTT TCCCAATAAA ATTGGTCGTT 900
 CTTGGTTTTT TATTGGAGA GTCTGTGCAA AATGTCATA AAAATAAATT AGCACTAGAA 960
 ATTATTTCTA AATACCAA

SEQ ID NO:244 PBQ8 DNA SEQUENCE

Nucleic Acid Accession#: X51405

Coding sequence: 3-1721 (underlined sequence corresponds to start and stop codon)

1 11 21 31 41 51

5
 10
 15
 20
 25
 30
 35
 40

```

AAATGGCGTG | CCCGTCCTCT | CGCCGGCCCC | CTGCCTCGCA | GTGGTTTCTC | CTGCAGCTCC 60
CCTGGGCTCC | GCGGCCAGTA | GTGCAGCCCC | TGGAGCCGCG | GCTTTGCCCG | TCTCCTCTGG 120
GTGGCCCCAG | TGCCCGGGCT | GACACTCATT | CAGCCGGGGA | AGGTGAGGCG | AGTAGAGGCT 180
GGTGCAGAAC | TTGCCGCCCC | CAGCAGCGCC | GCGGGGCTAA | GCCAGGGGCC | GGGCAGACAA 240
AAGAGGCCGC | CCGCGTAGGA | AGGCACGGCC | GCGGGCGGCG | GAGCGCAGCG | ATGCGCGGGC 300
GAGGGGGCAG | CCGCGTGCTG | GCTCTGTGCG | GGGCACTGGC | TGCCCTGCGG | TGGCTCCTGG 360
GCGCCGAAGC | CCAGAGCCCC | GGGGCGCCCG | CGGCGGGCAT | GAGGCGGCGC | CGGCGGCTGC 420
AGCAAGAGGA | CGGCATCTCC | TTCAGTACC | ACCGCTACCC | CGAGCTGCGC | GAGGCGCTCG 480
TGTCCTGTGT | GCTGCAGTGC | ACCGCCATCA | GCAGGATTFA | CACGGTGGGG | CGCAGCTTCG 540
AGGCGCGGGA | GCTCCTGGTC | ATCGAGCTGT | CCGACAACCC | TGGCGTCCAT | GAGCCTGGTG 600
AGCCTGAATT | TAAATACATT | GGGAAATATG | ATGGGAATGA | GGCTGTTGGA | CGAGAACTGC 660
TCATTTTCTT | GGCCAGTAC | CTATGCAACG | AATACCAGAA | GGGGAACGAG | ACAATTGTCA 720
ACCTGATCCA | CAGTACCCGC | ATTACATCA | TGCTTTCCCT | GAACCCAGAT | GGCTTTGAGA 780
AGGCAGCGTG | TCAGCCTGGT | GAACTCAAGG | ACTGGTTTGT | GGGTCGAAGC | AATGCCCAGG 840
GAATAGATCT | GAACCGGAAC | TTTCCAGACC | TGGATAGGAT | AGTGACGTG | AATGAGAAAG 900
AAGTGTGTCC | AAATAATCAT | CTGTGAAAA | ATATGAAGAA | AATTGTGGAT | CAAAACACAA 960
AGCTTGCTCC | TGAGACCAAG | GCTGTCAATC | ATTGGATTAT | GGATATTCCT | TTTGTGCTTT 1020
CTGCCAATCT | CCATGGAGGA | GACCTTGTGG | CCAATTATCC | ATATGATGAG | ACGCGGAGTG 1080
GTAGTGCTCA | CGAATACAGC | TCCTCCCCAG | ATGACGCCAT | TTTCCAAAGC | TTGGCCCGGG 1140
CATACTCTTC | TTTCAACCCG | GCCATGTCTG | ACCCAATCG | GCCACCATGT | CGCAAGAAATG 1200
ATGATGACAG | CAGCTTTGTA | GATGGAACCA | CCAACGGTGG | TGCTTGGTAC | AGCGTACCTG 1260
GAGGGATGCA | AGACTTCAAT | TACCTTAGCA | GCAACTGTTT | TGAGATCACC | GTGGAGCTTA 1320
GCTGTGAGAA | GTTCCACCT | GAAGAGACTC | TGAAGACCTA | CTGGGAGGAT | AACAAAAACT 1380
CCCTCATTAG | CTACCTTGAG | CAGATACACC | GAGGAGTTAA | AGGATTTGTC | CGAGACCTTC 1440
AAGSTAACCC | AATTGCGAAT | GCCACCATCT | CCGTGGAAGG | AATAGACCAC | GATGTTACAT 1500
CCGCAAGGA | TGGTGATTAC | TGGAGATTGC | TTATACCTGG | AAACATATAA | CTTACAGCCT 1560
CAGCTCCAGG | CTATCTGGCA | ATAACAAAGA | AAGTGGCAGT | TCCTTACAGC | CCTGCTGCTG 1620
GGGTGTGATT | TGAACGTGG | TCATTTTCTG | AAAGGAAAGA | AGAGGAGAAG | GAAAGAAATGA 1680
TGGAAATGGT | GAAATATGAT | TCAGAAACTT | TAAATTTTAA | AAAAGGCTTC | TAGTTAGCTG 1740
CTTTAAATCT | ATCTATATAA | TGTAGTATGA | TGTAAATGTT | TCTTTTTTT | AGATTTTGTG 1800
CAGTTAATAC | TTAACATTGA | TTTATTTT | AATCATTTAA | ATATTATCA | ACTTTCCCTTA 1860
AAATAAATAG | CCTCTTAGGT | AAAAATATAA | GAACCTTGATA | TATTTTCATTC | TCTTATATAG 1920
TATTCATTTT | CCTACCTATA | TTACACAAAA | AAGTATAGAA | AAGATTTAAG | TAAATTTGCC 1980
ATCCTAGGCT | TAAATGCAAT | ATTCCTGGTA | TTATTTACAA | TGCAGAAATT | TTTGAGTAAT 2040
TCTAGCTTTT | AAAAATTAGT | GAACTTCTTT | TACTGTAATT | GGTGACAATG | TCACATAATG 2100
AATGCTATTG | AAAAGGTTAA | CAGATACAGC | TCGGAGTTGT | GAGCACTCTA | CTGCAAGACT 2160
TAAATAGTTC | AGTATAAAT | GTCTGTTTTT | TCTTGTGCTG | ACTAATCTATA | AGCATGATCT 2220
TGTTAATGCA | TTTTGTGATG | GAAGAAAAGG | TACATGTTTA | CAAAGAGGTT | TTATGAAAAG 2280
AATAAAATTT | GACTTCTTGC | TTGTACATAT | AGGAGCAATA | CTATTATATT | ATGTAGTCCG 2340
TTAACTACTAC | TTTAAAGTTT | AGGGTTTTCT | CTGTTGTTGA | GAGTGGCCCA | GAATTGCATT 2400
CTGAATGAAT | AAAGGTTTAA | AAAAATCC | CAGTGAAAAA | AAA
  
```

SEQ ID NO:245 PBQ8 Protein sequence

Protein Accession#: P16870

5
 10
 15
 20
 25
 30
 35
 40
 45

```

MAGRGGSSALL ALCGALAACG WLLGAEAQEP GAPAAGMRRR RRLQEDGIS FEYHRYPELR 60
EALVSVWLQC TAISRIYTVG RSFEGRELLV IELSDNPGVH EPGEPEFKYI GNMHGNEAVG 120
RELLIFLAQY LCNEYQKQNE TIVNLHSTR IHIMPSLNPD GFKAASQPG ELKDWVFGRS 180
NAQGDILNRN FPDLDRIYV NEKEGGPNNH LLKNMKKIVD QNTKLAPETK AVIHWMIDIP 240
FVLSANLHGG DLVANYPYDE TRSGSAHEYS SSPDDAIFQS LARAYSSFN AMSDPNRPCC 300
RKNDDSSSFV DGTINGGAWY SVPGGMQDFN YLSSNCFEIT VELSCKEFP EETLKTYWED 360
NKNSLISYLE QIHRGVKGFV RDLQGNPIAN ATISVEGIDH DVTSAKDGDY WRLLIPGNYK 420
LTASAPGYLA ITKVAVPYS PAAGVDFELE SFSEKKEEEK EELMEWWKMM SETLNF
  
```

SEQ ID NO:246 PB4 DNA sequence

Nucleic Acid Accession#: AF038966

Coding sequence:

91-1107 (underlined sequence corresponds to start and stop codon)

60
 65
 70
 75
 80

```

1 11 21 31 41 51
GGGGCGACGT | GAGCGCGCAG | GGGGGCGGCG | GCCTCGCCTC | GTCTCTCTCT | CTGCGCCTGG 60
GTCGGGTGGG | TGACCCGAG | AGCCAGAGAG | ATGTCGGATT | TCGACAGTAA | CCCGTTTGCC 120
GACCCGGATC | TCAACAATCC | CTTCAGGAT | CCATCAGTTA | CACAAGTGAC | AAGAAATGTT 180
CCACCGGAC | TTGATGAATA | TAATCCATTC | TCGGATTTCTA | GAACACCTCC | ACCAGGCGGT 240
GTGAAGATGC | CTAATGTACC | CAATACACAA | CCAGCAATAA | TGAACCAAC | AGAGGAACAT 300
CCAGCTTATA | CACAGATTGC | AAAGGAACAT | GCATTGGCCC | AAGCTGAACT | TCTTAAGCGC 360
CAAGAAGAAC | TAGAAGAAA | AGCCGCAGAA | TTAGATCGTC | GGAACGAGA | AATGCAAAAC 420
CTCAGTCAAC | ATGGTAGAAA | AAATATTGG | CCACCTCTTC | CTAGCAATTT | TCCTGTCCGA 480
CCTTGTCTCT | ATCAGGAATT | TTCTGTAGAC | ATTCTGTAG | AATTCCAAA | GACAGTAAAG 540
CTTATGTAAT | ACTTGTGGAT | GTTCCATGCA | GTAACACTGT | TTCTAAATAT | CTTCGGATGC 600
TTGGCTTGGT | TTTGTGTTGA | TTCTGCAAGA | GCGGTGATT | TTGGAATTGAG | TATCCTGTGG 660
TTCTTGCTTT | TTACTCCTTG | TTCTTTGTC | TGTGTTGACA | GACCACTTTA | TGGAGCTTTC 720
AGGAGTGACA | GTTCATTAG | ATCTTTGTA | TTCTTCTCG | TCTATATTG | TCAGTTTGCT 780
GTACATGTAC | ATCAGGATTC | AGGATTTTCA | AACTGGGGCA | ATTGTGGTTG | GATTTTCATCC 840
CTTACTGGTC | TCAACCAAAA | TATTCCTGTT | GGAATCATGA | TGATAATCAT | AGCAGCACTT 900
TTACAGCAT | CAGCAGTCT | CTCAGTATT | ATGTTCAAAA | AAGTACATGG | ACTATATCGC 960
ACAACAGGTG | CTAGTTTGA | GAAGGCCCAA | CAGGAGTTTG | CAACAGGTGT | GATGTCCAAC 1020
AAAACGTGCC | AGACCGCAGC | TGCAAAATGCA | GCTTCAACTG | CAGCATCTAG | TGCAGCTCAG 1080
  
```

5 AATGCTTTCA AGGGTAACCA GATTTAAGAA TCTTCAAACA ATACACTGTT ACCTTTTGAC 1140
 TGTACCTTTT TCTCCAGTTA CTGTATTCFA CAAATATTTT TATGTTCAAA ACACACAGTA 1200
 CAGACAGCAT GGATATTTCC TGTTCACCTTG TGCATGGGCT AAAACCAGGA AAACCTTCCTT 1260
 GTCTTATTAC TTACCTTAAT AGTTTCTTAA TATTTTCAGTG CCCCTTGCAG AAAAAATATT 1320
 ACATGCTAAA TAAATATCTT CCATATTTT TGGGGATGAC ATTTCAGTGAA TTATTTTCAGT 1380
 GGTGACCCAC TGAAAATTAA TAATGGTACT TATGATTAAA AACGCATTTA ATACTAAGTG 1440
 CAGTAGTTCT TTCAAGAATC TTTAGAGATA AGGATTGCAC ATTGGAAAAG TAAACCATGT 1500
 TTCATTCTCT TTTCCCTATT TATATTGAAA GAAATAGGCC AGCAGAGACT TAGGGATTTT 1560
 10 AAATGGCTT GCTTTTAGC TGTTCAGTC ACCAGTGAAG AGCCTATGTG CATTTTGTAG 1620
 TAGATAATGT AAAATTTGTC ATCTTTTCTT TTTCTTTT TTAGAATAGC TGATATTTTG 1680
 ATAACAATCT CTAATTTGCA TGGGCACCAC ATTTCTTATA TTAAAAGAA TAGTGTTTGT 1740
 GCTTCTGTAC TGCTTATGGT TGTAGGATTC AGGGGTAAAT GGAATCACAG AAATGATATT 1800
 CTGCAAGAAT TTCTTTTAAA TAAAAAGTTT GGGGGTGCAA TATAAGAAAT TTATATAATA 1860
 15 TGCAGTACAT TATCCAAAAG AGAAGTAGT TAATGCAGTA GAAAGTAGTG GTAATAATTTC 1920
 CTTTTT

SEQ ID NO: 247 PB4 Protein sequence:

Protein Accession #:

20 MSDFDSNPFA DPDLNNPFKD PSVTQVTRNV PPGLDEYNPF SDSRTPPPGG VKMPNVPNTQ 60
 PAIMKPTEEH PAYTQIAKEH ALAQAEELKR QEELERKAAE LDRREREMQN LSQHGRKNIW 120
 PPLPSNFPVG PCFYQEFSDV IPVEFQKTVK LMYLWMFHA VTLFLNIFGC LAWFCVDSAR 180
 25 AVDFGLSILW FLLTPCSFV CWYRPLYGAF RSDSSFRFFV FFFVYICQFA VHLQAAGFH 240
 NWGNCGWISS LTGLNQNPV GIMMILAAAL FTASAVISLV MFKKVHGLYR TTGASFKAQ 300
 QEFATGVMSN KTVQTAAANA ASTAASSAAQ NAFKGNQI

SEQ ID NO:248 PBH2 DNA sequence

Nucleic Acid Accession#:

Coding sequence: 1-613 (underlined sequence corresponds to start and stop codon)

35 ATGAGAGACA ATAAATCGTG TGCTTTTTTC ATGGGAAAGT TAAATGTTTG TTTTGAAGGC 60
 ACAGTAATAG CAGGCTATTC AGTGTGTCCT ACTACCTGCA TCATTCATCT GGCTGTAGCT 120
 AGTGCACTAC AATTTCTTAA AAAGTCTTCT CACCCTCACA GGACTGCTCT ACATCTGGCC 180
 TCTGCCAATG GAAATTCAGA AGTAGTAAAA CTCCTGCTGG ACAGACGATG TCAACTTAAT 240
 ATCCTTGACA ACAAAAAGAG GACAGCTCTG ACAAAGGCCG TACAATGCCA GGAAGATGAA 300
 40 TGTGCGTTAA TGTGCTGGA ACATGGCACT GATCCGAATA TTCCAGATGA GTATGGAAAT 360
 ACCGCTCTAC ACTATGCTAT CTACAATGAA GATAAATTAA TGGCCAAAGC ACTGCTCTTA 420
 TACGGTGCTG ATATCGAATC AAAAAACAAG CATGGCCTCA CACCACTGTT ACTTGGTGTA 480
 CATGAGCAAA AACAGCAAGT GGTGAAATTT TTAATCAAGA AAAAAAGCAA TTAAATGCA 540
 CTGGATAGAT ATGGAAGGTG TGTGACCTTG GGAACGTTAT TTACCACCAA ATATGTTGTC 600
 45 ATATATGAAA AGTAG

SEQ ID NO:249 PBH2 Protein sequence:

Protein Accession #:

none found

50 MRDNKSCAFF MGKLVNCFEG TVIAGYSVFA TTCIHLAVA SALQFPKSS HPHRTALHLA 60
 SANGNSEVVK LLLDRRCQLN ILDNKKRTAL TKAVQCQEDE CALMLLEHGT DPNIPEYGN 120
 TALHYAIYNE DKLMAKALL YGADIESKNK HGLTPLLLGV HEQKQVVKF LIKKKANLNA 180
 LDRYGRCVTL GTLFTTKYVV IYK

SEQ ID NO:250 PB1 DNA sequence

Nucleic Acid Accession#:

Coding sequence: 1-3043 (underlined sequence corresponds to start and stop codon)

60 ATGGTGATCA TCTATCTTTC TTCTGCAAT TATTACATGG AGTTCTACAG AGAAGAGCTT 60
 CCCACATTG ACTATTTGAT TGACATTCAG TTTGCAACAG GAAAGGTTAC TCAGCCGGGA 120
 GAGGACACTT CCTACATCA ATGCGCTCAG CTGGAAGCCA GAGACGAAGG CACCGACAGT 180
 65 TTATTATTAA ACAATGGCAG CAGCGCCACG CTGAAGACAC GAACGCGCTG TTATGGAACC 240
 CCCAGAGGTC TCCCCATCG TAGCCTGCTC CAGCCGACTC CGCCACATG TAAAACGAAG 300
 ATCAGGAGCA GATTGAAGA ATTACAAAGT GAATTGGTGC CAGTCAGCAT GTCAGAGACA 360
 GACCACATAG CCTCTACTTC CTCTGATAAA AATGTTGGGA AAACACCTGA ATTAAAGGAA 420
 GACTCATGCA ACTTGTTTTC TGGCAATGAA AGCAGCAAAT TAGAAATGA GTCCAAACTA 480
 TTGTCATTAA AACTGTATAA AACTTTATGT CAACCTAATG AGCATAATA TCGAATTGAA 540
 70 GCCCAGGAAA ATTATATTC AGATCATGGT GGAGGTGAGG ATTCTGTGTC CAAAACAGAC 600
 ACAGGCTCAG AAAATCTGTA ACAAATAGCT AATTTTCTTA GTGGAATTT TGCTAAACAT 660
 ATTTCAAAAA CAAATGAAAC AGAACAGAAA GTAACACAAA TATTGGTGGA ATTAAGGTCA 720
 TCTACATTTC CAGAATCAGC TAATGAAAAG ACTTATTCAG AAAGCCCTTA TGATACAGAC 780
 TGCACCAAGA AATTATTATC AAAAAATAAG AGCGTTTCAG CATCAGAGGA TTTGTTGGAA 840
 75 GAAATAGAA CTGAGCTCTT ATCTCGGAG TTTGCAGAAC ATCGAGTACC AAATGGAATG 900
 AATAAGGGAG AACATGCATT AGTTCTGTTT GAAAAGTGTG TGCAAGATAA ATATTGCAG 960
 CAGGAACATA TCATAAAAAA GTTAATTAAG GAAAATAAGA AGCATCAGGA GCTCTTCGTA 1020
 GACATTGTTT GAGAAAAAGA CAATTTAAGA GAAGAATAAG AGAAAAGAAC AGAAACTGAG 1080
 AAGCAGCATA TGAACACAAT TAAACAGTTA GAATCAAGAA TAGAAGAACT TAATAAAGAA 1140
 GTTAAAGCTT CCAGAGATCA ACTAATAGCT CAAGACGTTA CAGTAAAAA TGCAGTTTCA 1200

CAGTTACACA AAGAGATGGC CCAACGGATG GAACAGGCCA ACAAGAAATG TGAAGAGGCA 1260
 CGCCAAAGAAA AAGAAAGCAAT GGTAAATGAAA TATGTAAGAG GTGAGAAGGA ATCTTTAGAT 1320
 CTTCGAAAGG AAAAAGAGAC ACTTGAGAAA AAACCTAGAG ATGCAAATAA GGAACCTGAG 1380
 5 AAAAACAATA ACAAATTAAC GCAGCTTCT CAGGAGAAAG GACGGTTGCA CCAGCTGTAT 1440
 GAAACTAAGG AAGGCGAAAC GACTAGACTC ATCAGAGAAA TAGACAAATT AAAGGAAGAC 1500
 ATTAACCTC ACGTCATCAA AGTAAAGTGG GCACAAAACA AATTAAAAGC TGAATGGAT 1560
 10 TCACACAAGG AAACCAAAGA TAAACTCAA GAAACAACAA CAAAAATTAAC ACAAGCAAAG 1620
 GAAGAAGCAG ATCAGATACG AAAAAACTGT CAGGATATGA TAAAAACATA TCAGGAGTCA 1680
 GAAGAAATTA AATCAATGA GCTTGATGCA AAGCTTAGAG TCACAAAAGG AGAAGCTTGA 1740
 AAACAAATGC AAGAAAAATC TGACCAGCTA GAGATGCATC ATGCCAAAT AAAGGAACTA 1800
 GAAGATCTGA AGAGAACATT TAAGGAGGGT ATGGATGAGT TAAGAACACT GAGAACAAAG 1860
 GTGAAATGTC TAGAAGATGA ACGATTAAGA ACAGAAGATG AATTATCAA ATATAAGGAA 1920
 ATTATTAATC GCCAAAAAGC TGAATTCAG AATTATTGG ACAAGGTGAA AACTGCAGAT 1980
 15 CAGCTACAGG AGCAGCTTCA AAGAGGTAAG CAAGAAATTG AAAATTTGAA AGAAGAAGTG 2040
 GAAAGTCTTA ATCTTTGAT TAATGACCTA CAAAAAGACA TCGAAGGCAG TAGGAAAAAG 2100
 GAATCTGAGC TGCTGCTGTT TACAGAAAGG CTCACTAGTA AGAATGCACA GCTTCAGTCT 2160
 GAATCCAATT CTTGTCAGTC ACAATTTGAT AAAGTTTCT GTAGTGAAAG TCAGTTACAA 2220
 AGCCAGTGTG ACAAATGAA ACAGACAAAT ATTAATTTGG AAAGTAGGTT GTTGAAAGAG 2280
 20 GAAGAACTGC GAAAAAGAGG AGTCCAAACT CTGCAAGCTG AACTCGCTG TAGACAAACA 2340
 GAAGTTAAAG CATTGAGTAC CCAGGTAGAA GAATTAAGG ATGAGTTAGT AACTCAGAGA 2400
 CGTAAACATG CCTCTAGTAT CAAGGATCTC ACCAAACAA TCAGCAAGC ACGAAGAAAA 2460
 TTAGATCAGG TTGAGAGTGG AAGCTATGAC AAAGAAGTCA GCAGCATGGG AAGTCGTTCT 2520
 AGTTCCAGG GGTCCCTGAA TGCTCGAAGC AGTGCAGAAG ATCGATCTCC AGAAAAACT 2580
 25 GGGTCTCAG TACTGTGGA TAACTTTCCA CAAGTAGATA AGGCCATGTT GATTGAGAGA 2640
 ATAGTTAGGC TCACAAAAGC ACATGCCCCG AAAAAATGAA AGATAGAATT TATGGAGGAC 2700
 CACATCAAACT AACTGGTGGG AGAAATTAGG AAAAAACAA AAATAATTCA AAGTTATATT 2760
 TTACGAGAAG AATCAGGCAC ACTTCTTCCA GAGGCATCTG ATTTTAAACA AGTTCATTTA 2820
 AGTAGACGGG GTGGCATCAT GGCATCTTTA TATACATCCC ATCCAGCTGA CAATGGATTA 2880
 30 ACATTGGAGC TCTCTTGGG AATCAACCGA AAATTACAGG CTGTTTGGG GGATACGTTA 2940
 CTAAAAAATA TTACTTTGAA GGAAATCTA CAAACACTTG GAACAGAAAT AGAACGTCCT 3000
 ATTAACACACC AGCATGAATC AGAACAGAGG ACAAAGAAAA CCTAAAACA GCCTCTTGCT 3060
 CAGTAAAGAG ACAAAGGCCA CACAGGAGTA GGTGCCACTG ACCTCTATTG TTGGAGACTT 3120
 TGTTCCACTT TTGTTCAG CCAGTAAAAA TATGTTTTG CTCATCTGT ACACAAAAA 3180
 35 ATACCTTTT ACAATATGAA TGCAATGCTG TATATACTGT AAGACTGAAA GCCTTGATGA 3240
 AATTGTGTTT TGTATGGTGC AATATGACAG CCGTCATTG AATCTAAACA ACTTAATTG 3300
 CTTGTATTCA TAAGAAGTGT TGAACATTAC AAGGGCTTTT AT

Protein Accession #: SEQ ID NO:251 PBj1 Protein sequence:
 NP_060487

MVIIYLSFCN YYMEFYREEL PHIDYLDIQ FATGKVTPG EDTSYHQCAQ LEARDEGTD 60
 LLLNNGSSAT LKTRTRCYGT PRGLPHRSL QPTPTCKTK IRSRFEELQS ELVPVMSSET 120
 45 DHIASSTSDK NVGKTPELKE DSCNLFSGNE SSKLENESKL LSLNTDKTLC QPNEHNNRIE 180
 AQENYIPDHG GGEDSCKATD TGSENSEIA NFPSGNFAKH ISKTNETEQK VTQILVELRS 240
 STIPESANEK TYSESPYDID CTKKFISKIK SVSASEDLLE EIESELLSTE FAHRVPMNGM 300
 NKGEHALVLF EKCVDQKYLQ QEHIIKKLIK ENKKHQELFV DICSEKDNLR EELKKRTETE 360
 KQHMNTIKQL ESRIEELNKE VKASRDQLIA QDVTAKNVQ QLHKEMAQRM EQANKKCEEA 420
 50 RQKEAMVMK YVRGEKESLD LRKEKETLEK KLRDANKLE KNTNKKIKLS QEKGRHLQLY 480
 ETKEGETIRL IREIDKLKED INSHVIKVKW AQNKLKAEMD SHKETKDKLK ETITKLQAK 540
 EADQIRKNC QDMIKTYQES EEIKSNELDA KLRVTKGELE KQMQEKSQDL EMHHAKEIKEL 600
 EDLKRFTKEG MDELRLTRTK VKCLEDERLR TEDELSKYKE INROKAEIQ NLLDKVKTAD 660
 QLQEQLRQG QEIEENLKEEV ELSNLSLNDL QKDIEGSRKR ESELLFTER LTSKNAQLQS 720
 55 ESNLSQSQFD KVSCEQMLQ SQCEQMKQTN INLESRLKE EELRKEEVQT LQAEALACROT 780
 EVKALSTQVE ELKDELVTQR RKHASSIKDL TKQLQARRK LDQVESGSYD KEVSSMGSR 840
 SSSGSLNARS SAEDRSPTNT GSSVAVDNFP QVDKAMLIER IVRLQKAHAR KNEKIEFMD 900
 HIKQLVEIIR KTKIIQSYI LREESGLSS EASDFNKVHL SRRGGIMASL YTSHPADNGL 960
 TLELSLEINR KLQAVLEDITL LKNITLKENL QTLGTEIERL IKHQHELEQR TKKT

SEQ ID NO:252 PBj6 DNA sequence

Nucleic Acid Accession#: D83760
 Coding sequence: 56-1459 (underlined sequence corresponds to start and stop codon)

1	11	21	31	41	51	
TTGCCGTGAA	GGGCTGTGCG	GTTCCCGTGC	GCGCCGGAGC	CTGCTGTGGC	CTCTTATGCA	60
CTCCACCACC	CCCATCAGCT	CCCTCTCTC	CTTCACCAGC	CCCGCAGTGA	AGAGACTGCT	120
AGGCTGGAAG	CAAGGAGATG	AAGAGGAAAA	GTGGGCAGAG	AAGGCAGTGG	ACTCTCTAGT	180
GAAGAAGTTA	AAGAAGAAGA	AGGGAGCCAT	GGACGAGCTG	GAGAGGGCTC	TCAGCTGCCC	240
GGGGCAGCCC	AGCAAATGCG	TCACGATTCC	CCGCTCCCTG	GACGGGCGGC	TGCAGGTGTC	300
CCACCGCAAG	GGCCTGCCCC	ATGTGATTTA	CTGTCGCGTG	TGGCGCTGGC	CGGATCTGCA	360
GTCCCAACAC	GAGCTGAAGC	CGCTGGAGTG	CTGTGAGTTC	CCATTGGGCT	CCAAGCAGAA	420
AGAAGTGTGC	ATTAACCCCT	ACCACATACG	CCGGGTGGAG	ACTCCAGTAC	TGCCTCCTGT	480
GCTCTGCGCA	AGACACAGTG	AATATAACCC	CCAGCTCAGC	CTCCTGGCCA	AGTTCCGCG	540
CGCCTCCCTG	CACAGTGAGC	CACATCATGC	ACACAACGCC	ACCTATCTTG	ACTCTTTCCA	600
GCAGCTCCCG	TGCTCTGCAC	TCCCTCCCTC	ACCCAGCCAC	GCGTTCTCCC	AGTCCCCGTG	660
CACGGCCAGC	TACCTCTACT	CCCCAGGAAG	TCCTTCTGAG	CCAGAGAGTC	CCTATCAACA	720
CTCAGTTGAC	ACACCACCCC	TGCCCTATCA	TGCCACAGAA	GCCTCTGAGA	CCCAGAGTGG	780

5 CCAACCTGTA GATGCCACAG CTGATAGACA TGTAGTGCTA TCGATACCAA ATGGAGACTT 840
 TCGACCAGTT TGTACGAGG AGCCCCAGCA CTGGTGCTCG GTGCGCTACT ATGAACTGAA 900
 CAACCGAGTT GGGGAGACAT TCCAGGCTTC CTCCCGAAGT GTGCTCATAG ATGGGTTCAC 960
 CGACCCCTCA AATAACAGGA ACAGATTCTG TCTTGGACTT CTTTCTAATG TAAACAGAAA 1020
 CTCAACGATA GAAATACCA GGAGACATAT AGGAAAGGGT GTGCACTTGT ACTACGTCGG 1080
 GGGAGAGGTG TATGCCGAGT GCGTGAGTGA CAGCAGCATC TTTGTGCAGA GCCGGAACATG 1140
 CAACATATCA CACGGCTTCC ACCCAGCTAC CGTCTGCAAG ATCCCCAGCG GCTGCAGCCT 1200
 CAAGGTCTTC AACAAACAGC TCTTCGCTCA GCTCCTGGCC CAGTCAGTTC ACCACGGCTT 1260
 10 TGAAGTCGTG TATGAACCTGA CCAAGATGTG TACTATCCGG ATGAGTTTGT TTAAGGGTTG 1320
 GGGTGCTGAG TATCATCGCC AGGATGTCAC CAGCACCCCC TGCTGGATTG AGATTTCATCT 1380
 TCATGGGCCA CTGCAGTGGC TGGACAAAGT TCTGACTCAG ATGGGCTCTC CACATAACCC 1440
 CATTTCTTCA GTGTCTTAAC AGTCATGTCT TAAGCTGCAT TTCCATAGGA T

15 **SEQ ID NO:253 PBj6 Protein sequence:**
 Protein Accession #: NP_005896
 MHSTTPISSL FSFTSPAVKR LLGWKQGDDE EKWAIEKAVDS LVKKLKKKKG AMDELERALS 60
 20 CPGQPSKCVT IPRSLDGRQL VSHRKGPHV IYCRVWRWPD LQSHHEKPL ECCEFFPGSK 120
 QKEVCINPYH YRRVETPVL PVLVPRHSEY NPQLSLLAKF RSASLHSEPL MPHNAITYPDS 180
 FQQPSCSALP PPSHAFSOS PCTASYPHSP GSPSEPEPY QHSVDTPPLP YHATEASETQ 240
 SGQPV DATAD RHVVL SIPNG DFRPVCYEEP QHWCSVAYYE LNNRVGETFQ ASSRSVLIDG 300
 25 FTDPSSNNRRN FCLGLLSNVN RNTIENTRR HIGKGVHLYY VGGEVYAEVC SDSSIFVQSR 360
 NCNYQHGFHP ATVCCKIPSGC SLKVFNNQLF AQLLAQSVHH GFEVYVELTK MCTIRMSFVK 420
 GWGAIEYHRQD VTSTPCWIEI HLHGPLQWLD KVLTMGSPH NPISSVS

30 **SEQ ID NO:254 PBj8 DNA sequence**
 Nucleic Acid Accession#: AB04684
 Coding sequence: 472-4377 (underlined sequence corresponds to start and stop codon)

35 1 TGCAGGTTTG CAGGGTCTGA GATTACTTGG GCTTTTCCTG CTTTTCCTT TTGCTTAAGG 60
 GATGGACAAG GAGCTGAGAT TTATGACCCCT TATTAGAGAA AAAAATGTGC CTGTCTAGGG 120
 TGGGGACACT TGGTTGATGC AGTCTCTCTC TCTCTTCTC GGTGTTTATA ACAAACAAA 180
 ACCAAAATGA ACTGAGGGGT TTGTAATGGT AGTTTGTTTG TTGCTGGAGA ATGCTACTTT 240
 GCATGCTTTT TTCTCTTTCG AGGGTATGTT CTGCTTGTG CTTTTCCTT TAGAAGCTAC 300
 40 TAAAGGGTGT TGGGGATGCT TCTGACTATT ATGAAGGCCA AAAGGCTGT TGACTGGGGC 360
 TGCTTTTAAC CCTTCTCTAT TTGCTGAGAA TGCAGCCGTG TGACAGTAAC TGAACATTGG 420
 TCTAAGTCTT TTCCAAAAGG TCAAGGTTCA CAAGAACATC TGCTCAAAAT AATGACCATG 480
 GGGGATATGA AGACCCAGA CTTTGATGAC CTCTGGCAG CATTTGACAT CCCAGATATG 540
 GTCGATCCTA AAGCAGCTAT TGAGTCTGGA CACGATGACC ATGAAGGCCA CATGAAGCAG 600
 45 AATGCTCAGG GAGAGGATGA CTCACGCA CCATCATCTT CTGATGTGGG TGTGAGCGTT 660
 ATCGTCAAGA ATGTTGCGAA CATTGACTCT TCCGAGGGCG GGGAGAAAGA CGGCCACAAC 720
 CCCACTGGCA ATGGCTTACA TAATGGGTTT CTCACAGCAT CCTCCCTTGA CAGTTACAGT 780
 AAAGATGGAG CAAAGTCTT GAAAGGAGAT GTGCCTGCCT CTGAGGTGAC ACTGAAAGAC 840
 TGCAGATTCA GCCAGTTTAG CCCGATCTCC AGTGTGTAAG AGTTTGTATGA CGACGAGAAG 900
 50 ATTGAGGTGG ATGACCCCTC TGACAAGGAG GACATGCGAT CAAGCTTCAG GTCGAATGTG 960
 TTGACGGGGT CGGCTCCCCA GCAGGACTAC GATAAGCTGA AGGCCTCGG AGGGGAAAAC 1020
 TCCAGCAAAA CTGGACTCTC TACGTGAGG AATGTGGAGA AAAACAAAGC TGTTAAGAGA 1080
 GAAACAGAAG CAGATTCTAT AAACCTGAGT GTTTATGAAC CTTTAAAGT CAGAAAAGCA 1140
 GAGGATAAAT TGAAGGAAAG CTCTGACAAG GTGCTGGAAG ACAGAGTCTT AGATGGGAG 1200
 55 CTGAGCTCCG AGAAGAAATG CACCAGCCTC CCCAGCGTTG CGCCATCAA GACAAAGTCG 1260
 TCCTCCAAGC TCTCGTCTG CATCGCTGCC ATCGCGGCTC TCAGCGCTAA AAAGCGGCT 1320
 TCAGACTCCT GCAAAGAACC AGTGGCCAAT TCGAGGGAAT CCTCCCCGTT ACCAAAAGAA 1380
 GTAAATGACA GTCCGAGAGC CGCTGACAAG TCTCCTGAAT CCCAGAATCT CATCGACGGG 1440
 60 ACCAAAAAAC CATCCCTGAA GCAACCGGAT AGTCCAGAA GCATCTCAAG TGAGAACAGC 1500
 AGCAAGGAT CCCCGTCTC TCCCGCAGG TCCACACCAG CAATCCCCAA AGTCCGCATA 1560
 AAAACCATTA AGACATCTT TGGGGAAATC AAGAGAACAG TGACCAGGGT ATTGCCAGAA 1620
 GTGGATCTTG ACTCTGGAAA GAAACCTTCC GAGCAGACAG CGTCCGTGAT GGCCTCTGTG 1680
 ACATCCCTTC TGTCTGTGTC AGCATCAGCC GCCGTCTCTT CCTCTCCCC CAGGGCGCCT 1740
 CTCCAGTCTG CGGTCTGTGAC CAATGCAGTT TCCCCTGCAG AGCTCACCCC CAAACAGGTC 1800
 65 ACAATCAAGC CTGTGGCTAC TGCTTTCTCT CCAGTGTCTG CTGTGAAGAC GGCAGGATCC 1860
 CAAGTCATTA APTTGAAGCT CGCTAACAAC ACCACGGTGA AAGCCACGGT CATATCTGCT 1920
 GCCTCTGTCC AGAGTGCCAG CAGCGCCATC ATTAAAGCTG CCAACGCCAT CCAGCAGCAA 1980
 ACTGTCTGTG TGCCTGCTG CAGCCTGGCC AATGCCAAAC TCGTGCCAAA GACTGTGCAC 2040
 CTTGCCAAC CTTAACCTTTT GCCTCAGGGT GCCCAGGCCA CCTCTGAAC CCGCAAGTG 2100
 70 CTAACCAAC CTAAGCAACA AATAAGCAG GCAATAATCA ATGCAGCAGC CTCGCAACCC 2160
 CCCAAAAGG TGTCTCGAGT CCAGGTGGTG TCGTCTTGC AGAGTTCTGT GGTGGAAGCT 2220
 TTCAACAAGG TGTGAGCAG TGTCAATCCA GTCCCTGTTT ACATCCCAA CCTCAGTCTC 2280
 CCCGCCAATG CAGGGATCAC GTTACCAGC CGTGGGTACA AGTGTCTTGA GTGTGGGGAC 2340
 75 TCCTTTGCAC TTGAAAAGAG TCTGACCCAG CACTACGACA GACGGAGCGT GCGCATCGAA 2400
 GTAACGTGCA ACCATTGTAT AAAGAACCCT GTTTTGTACA ACAAATGCAG CCTCCTTCC 2460
 CATGCCCGTG GGCATTAAGG GAAAGGGGTG GTAATGCAAT GCTCCCACTT AATTTTAAAG 2520
 CCACTCCAG CAGATCAAT GATAGTTTCT CCGTCAAGCA ATACTTCCAC TTCAACTTCC 2580
 ACTCTTCAGA GCCCTGTGGG AGCTGGCACA CACTGTGTC CAAAATTCAT GTCTGGCATA 2640
 ACTGGGACAG TCATATCAAGC TCTTCAAGC ACTCCCATCA CCCCAGCCAT GCCCTAGAT 2700
 80 GAAGACCCCT CCAAACTGTG TAGACATAGT CTAATAATGT TGGAGTGTAA TGAAGTCTTC 2760
 CAGGACGAGA CATCACTGGC TACACATTTC CAGCAGGCTG CAGATACGAG TGGACAAAAG 2820

5 ACTTGCACTA TCTGCCAGAT GCTGCTTCCT AACCAGTGCA GTTATGCATC ACACCAGAGA 2880
 ATCCATCAGC ACAAATCTCC CTACACCTGC CCTGAGTGTG GGGCCATCTG CAGGTCGGTG 2940
 CACTTCCAGA CCCACGTCAC CAAGAAGTGT CTGCACTACA CGAGGAGAGT TGGTTTTCGA 3000
 TGTGTGCATT GCAATGTTGT GTACTCTGAT GTGGCTGCTC TGAAGTCTCA CATTCAGAGT 3060
 TCTCACTGTG AAGTCTTCTA CAAGTGTCTT ATTTGTCCAA TGGCGTFTAA GTCTGCCCCA 3120
 AGCACACATT CCCACGCTTA CACACAGCAT CCTGGCATCA AGATAGGAGA ACCAAAAATA 3180
 ATATATAAGT GTTCCATGTG CGACACTGTG TTCACCTGTC AAACCTTGCT GTATCGCCAC 3240
 TTTGACCAAC ACATTGAAAA CCAGAAGGTG TCTGTTTCA AGTGTCAGAG CTGTTCTCTT 3300
 TTATATGCAC AGAAGCAACT TATGATGGAC CATATCAAGT CTATGCATGG AACATTGAAA 3360
 10 AGTATTGAAG GGCCTCCAAA CTTGGGTATA AACTTGCCCT TGAGCATTA GCCTGCAACT 3420
 CAAAATTGAG CAAATCAGAA CAAAGAGGAC ACCAAATCCA TGAATGGGAA AGAGAAATTG 3480
 GAAAAGAAAT CTCCATCTCC TGTGAAAAA TCAATGGAAA CCAAGAAAGT GGCCAGTCCT 3540
 GGGTGGACGT GTTGGGAGTG TGACTGCCTG TTCATGCAGA GAGATGTGTA CATATCCAC 3600
 GTGAGGAAAG AGCACGGGAA GCAAAATGAAG AAACACCCCT GCCGCCAGT TGACAAGTCT 3660
 15 TTCAGCTCGT CCCACGCGT GTGCCGCGAC AACCGGATCA AGCACAAAG CATCAGGAAA 3720
 GTGTACGCTT GCTCGCAGTG CCCAGACTCC AGACGTACCT TTACCAAACG TTTGATGCTG 3780
 GAGAAGCAGC TCCAGCTGAT GCATGGCATC AAGGACCTG ACCTGAAAGA AATGACAGAT 3840
 GCCACCAATG AGGAGGAAAC AGAAATAAAA GAAGACACTA AGGTCCCCAG TCCCAAGCGG 3900
 20 AAGTTGGAAG AACCAGTTCT GGAGTTCAGG CCTCCCCGAG GAGCAATCAC TCAACCACTG 3960
 AAAAAGCTGA AAATCAATGT TTTTAAGGTT CACAAGTGTG CCGTGTGTGG CTTCACCACC 4020
 GAAAACCTGC TGCAATTCCA CGAACACATC CCTCAGCACA AATCGGATGG TTCTTCTTAC 4080
 CAGTGCCGGG AGTGTGGCCT CTGCTACACG TCTCACGTCT CTCTGTCCAG GCACCTCTTC 4140
 ATCGTACACA AGTTAAAGGA ACCTCAGCCA GTGTCCAAGC AAAATGGGGC TGGGGAAGAT 4200
 25 AACCAACAGG AGAACAAACC GACCCAGCAG GATGAATCCC CTGATGGCGC CGTGTGAGAC 4260
 AGAAAGTGCA AAGTGTGCGC AAAAAGCTTTT GAAAGTGAAG CTGCCTTAAA TACTCACATG 4320
 CGGACACACG GCATGGCCCTT CATCAAAATCC AAAAGGATGA GCTCAGCCGA GAAATAGCCA 4380
 CAGATGCTCC ATGAGGAAAA TCCCTGTCCA CATTTGGAATA AAAAAGACAT TTTTGTATCA 4440
 AAGTTTGAGT TATAATAGAG TTAACAGTAC TGTCTAGGCT GTTGCAATAT ATTCTCTTTC 4500
 30 AATGTACCTT CCTTCACCTC GTCGTATATA TCCTCGATAA GTATTAAAC AGTATTGAG 4560
 TTTAAAGAG TTTGTATATA TTTAAATGAA TAACTTTTAA TACTCTTTGT TACATGTTTG 4620
 TATCAGTATT TAGTGGAAAA CCATTTGAGT TGTTTTGGGT TAGAATTTT CTTTTGTAC 4680
 TGTTCCTTTA AAACAGAGTT CTTAGTAACA GGGCAGTTC CTGAATTCAA ATAAACCAT 4740
 TTTGATGTTT GGATTTTGAA TGGGTTAACT AATTACAGGC TAAATAAATG CTTTTTTAG 4800
 35 TGTTTTAAAT TTTTGAATTT CACTACATAA ATTGTAAGTA ATGTGGGTC TCAAAAACAC 4860
 TAGGAACCTT TAAGTGCTT AGCACTTCCT CGATGTGCTT GCCCTGAGGG AGTGAGTTCA 4920
 CATTTGAGAC AACTGCACTC CAGTGTGGAC GTGCCCTTGT CTTCAGGCCA TGCCGAAGGG 4980
 TGTTTAAAGC AGTCTTGAGC GTCGCTCCTT TCCAGCCGT GGATAAAAAC TGAAGCTAGG 5040
 40 AATCTAATAA GGAATGCTGA TTTCTCAGT TCCATTTTGA GGAATGGGA AGGCTATTCT 5100
 AAAGAAAAAA ATGGGATTTG TTTTCTCGG AGATCTGCAA GGCTGGCTTT AAGAGCACAA 5160
 GGAGGGAAG TAACGAAAGG GCTGGACTAC TATAAAAGTT ACAAATACGT AGTTAGACCA 5220
 ATAGATTAT ATAGTCAGGT TTTTGTGATG TAATTTATTA ACTAATATTT ACAGAAACAC 5280
 AGCTAAGAA ATCAAGATTT TCTCTGGCTC TTGACAGAAA AAAATCAGTT GACTTAACCC 5340
 TTTGCTGTCA AAAGAGTTGG CGTTTCTGT TCTGGGTGCT ACTGCCAAAC GTTATGGTAC 5400
 45 TTAGAGTCGG GATGCACAA TTCAACCAAC GACTTATCAA TGCAGCCGCC TGTGTATTGC 5460
 AATTGGCCGT TACCTTAAGC ACTGAGCCAC CCGGGTTTGG TTCAGCCATTT TCAAGAAAGTA 5520
 TATTTAACGT CGGTAGTCTT GCTTTATTA AATGCAGCAG AGGTACTCTT CTGTCCCTTC 5580
 CGTTTATAGT TCTCTGAGAG AGTTCTATTT TTTGGTTTGG TTTTGTGTTT TCTTTTGCAT 5640
 TTTGTATCTT GTATTATCTT CTGAACATGT TTTGTACCTT TTTTTTT TTTTPTTTAA 5700
 50 GAAAGGAAT TCTTTTGTGT ATATATAGAT ACTTGCATGA TATACTGTAG TCAATGTTTC 5760
 GTTCTCAAA AGGTCTTGCT GCTGTGAGT GTTATGCACT CCATCCATCA TAACTGTATG 5820
 AAACACATTT CATATGTAAA TAAACGTGGG ACATTTG

SEQ ID NO:255 PB/J8 Protein sequence:
 BAB13455

MKTPDFDILL AAFDIPDMVD PKAAIESGHD DHESHMKQNA HGEDDSHAPS SSDVGVSVIV 60
 KNRNIDSSE GGEKDGHNPT GNGLHNGFLT ASSLDYSYKD GAKSLKGDVP ASEVTLKDS 120
 60 FSQFSPISSA EEFDDDEKIE VDDPPDKEDM RSSFRSNVLT GSAPQQDYDK LKALGGENSS 180
 KTGLSTSGNV EKNKAVKRET EASSINLSVY EPFKVRKAED KLESSDKVL ENRVLDGKLS 240
 SEKNDTSLPS VAPSKTKSSS KLSSCIAAIA ALSAKKAASD SCKEPVANSR ESSPLPKEVN 300
 DSPRAADKSP ESQNLIDGTK KPSSLKQPDSP RSISSENSK GSPSPAGST PAIPKVRIT 360
 IKTSSGEIKR TVTRVLPEVD LDSGKKPSEQ TASVMASVTS LLSPASAAV LSSPPRAPLQ 420
 65 SAVVNAVSP AELTPKQVTT KPVATAFLPV SAVKTAGSQV INKLANNIT VKATVISAAS 480
 VQSASSAIHK AANAQQQTV VVPASSLANA KLVPKTVHLA NLNLLPQGAQ ATSELRQVLT 540
 KPQQQIKQAI INAAASQPPK KVSRRVQVSS LQSSVVEAFN KVLSSVNPVP VYIPNLSPPA 600
 NAGITLPTRG YKCLECGDSF ALEKSLTQHY DRRSVRIEVT CNHCTKNLVE YNKCSLLSHA 660
 RGHKEKGVM QCSHLILKPV PADQMIVSPS SNTSTSTSL QSPVGAGTHT VTKIQSGITG 720
 70 TVISAPSTP ITPAMPLDED PSKLCRHSK CLECNEVFQD ETSLATHFQQ AADTSQGKTC 780
 TICQMLLPNQ CSYASHQRIH QHKSPTCPE CGAICRSVHF QTHVTKNCLH YTRRVGFRVC 840
 HCNVVYSDVA ALKSHIQGSH CEVFKCPC PMAFKSAPST HSHAYTQHPG IKIGEPKIIY 900
 KCSMCDTVFT LQTLTYRHFQ QHIENQKVSF FKPDCSLLY AQKQLMMDHI KSMHGTLSKI 960
 EGPENLGINL PLSIKPATQN SANQNKEDTK SMNGKEKLEK KSPSPVKKSM ETKKVASPGW 1020
 75 TCWECCLFM QRDVYISHV KEHGKQMKKH PCRQCDKFS SSHSLCRHNR IKHKGIRKVV 1080
 ACSHCPDSRR TFKRLMLEK HVQLMHGIKD PDLKEMTDAT NEEETEIKED TKVPSPKRKL 1140
 EEPVLEFRPP RGATQPLKK LKINVKVHK CAVCGFTTEN LLQFHEHIPQ HKSDGSSYQC 1200
 RECGLCYTSV VLSRHLFIV HKLKEPQVS QONGAGEDNQ QENKPSHEDE SPDGAVSDRK 1260
 CKVCAKTFET EALNTHMRT HGMAFIKSKR MSSAEK

SEQ ID NO:256 PBM1 DNA sequence

Nucleic Acid Accession#: AF111847

Coding sequence: 58-1608 (underlined sequence corresponds to start and stop codon)

5
10
15
20
25
30
35
40
45
50
55

```

1      11      21      31      41      51
|      |      |      |      |      |
TTTTCGTCTGA CTCTTACCGG TTGGCTGGGC CAGCTGCGCC GCGGCTCACA GCTGACGATG 60
GGGGACCCCA GCAAGCAGGA CATCTTGACC ATCTTCAAGC GCCTCCGCTC GGTGCCCACT 120
AACAAAGGTGT GTTTTGATTG TGGTGCCAAA AATCCCAGCT GGGCAAGCAT AACCTATGGA 180
GTGTTCCTTT GCATTGATTG CTCAGGGTCC CACCGGTCAC TTGGTGTTCA CTGAGTTTT 240
ATTTCGATCTA CAGAGTTGGA TTCCAACTGG TCATGGTTTC AGTTGCGATG CATGCAAGTC 300
GGAGGAAACG CTAGTGCATC TTCTTTT TTTT CATCAACATG GGTGTTCCAC CAATGACACC 360
AATGCCAAGT ACAACAGTCG TGCTGCTCAG CTCTATAGGG AGAAAAATCA ATCGCTCGCC 420
TCTCAAGCAA CACGGAAGCA TGGCACTGAT CTGTGGCTTG ATAGTTGTGT GGTTCACCT 480
TTGTCCCTC CACCAAAGGA GGAAGATTTT TTTGCCCTCT ACGTTTCTCC TGAGGTGAGT 540
GACACAGCGT GGGCATCAGC AATAGCAGAA CCATCTTCTT TAACATCAAG GCCTGTGGAA 600
ACCACTTTGG AATAAATGA AGGTGGACAA GAGCAAGGAC CRAAGTGTTGA AGGTCTTAAT 660
GTACCAACAA AGGCTACTTT AGAGGTATCC TCATATATAA AAAAGAAACC AAATCAAGCT 720
AAAAAAGGCC TTGGGGGCTC AAAAGGAAGT TTGGGAGCTC AGAACTGGC AAACACATGC 780
TTTAATGAAA TTGAAAAACA AGCTCAAGCT GCGGATAAAA TGAAGGAGCA GGAAGACCTG 840
GCCAAGGTGG TATCTAAAGA AGAATCAATT GTTTCATCAT TACGATTAGC CTATAAGGAT 900
CTTGAAATTC AATGAAGAA AGACGAAAAG ATGAACATTA GTGGCAAAAA AAATGTTGAC 960
TCAGACAGAC TCGCATGGG ATTTGGAAT TGCAGAAGTG TTTATTCACA TTCAGTGACT 1020
TCAGATATGC AGACCATAGA GCAGGAATCA CCCATTATGG CAAAACCAAG AAAAAAGTAT 1080
AATGATGACA GTGACGATTC ATATTTTACT TCCAGCTCAA GTTACTTTGA CGAGCCAGTG 1140
GAGTTAAGGA GCAGTTCTTT CTCTAGCTGG GATGACAGTT CAGATTCTTA TTGAAAAAAA 1200
GAGACCAGCA AAGTACTTGA AACAGTTCTG AAAACCACAG GCTATTTCAGA CAGACCTACT 1260
GCTCGCCGCA AGCCAGATTA TGAGCCAGTT GAAAATACAG ATGAGGCCCA GAAGAAGTTT 1320
GGCAATGTCA AGGCCATTTC ATCAGATATG TATTTTGGAA GACAATCCCA GGCTGATTAT 1380
GAGACCAGGG CCCGCCTAGA GAGGCTGTCT GCAAGTTCTT CCATAAGCTC GGCTGATCTG 1440
TTTCGAGGAGC CGAGGAAGCA GCCAGCAGGG AACTACAGCC TGTCCAGTGT GCTGCCCAAC 1500
GCCCCGACA TGGCGCAGTT CAAGCAGGGA GTGAGATCGG TTGCTGGAAA ACTCTCCGCT 1560
TTTGCTAATG GAGTCGTGAC TTCAATTCAG GATCGCTACG GTTCTTAATA CTGAAGTCAT 1620
GATGTGTATT TCCTGGAGAA ATTCCTCTTT AAATGAACAA GTAACCACAT CTCAGGCGGC 1680
AGTGAAGTCC AGATAGTTT GCAGATTGTT TTGCTACTTT TTCAATAGGT ATATGTTTCT 1740
GATTTTAAAT ATTTCTTTTG AGAAATCTCT AGTTCTGATG TAGGAGCTTT CCTGTGATT 1800
CTGTTTACG TTCTTCTCTG TCACACCCTC CTTTGGCGTC TCTGTGATA TCCTTGCTTT 1860
ATTTCTTTGG AACCTTTGAT TTCAACACTG AGGCGCTGGA GACCTCGGCT CCTCTGCTC 1920
CTGAACAGAG AGGCTTCATG TGGGGGAGGA GGAGAGGCTT CCATGTGACA CATGGGCTCA 1980
GGGCTGCCAG AATCAGCGGA TGTGTGATGG GCCTGCAGAA ACAACACTCA CCACACACAC 2040
TTCTTCAAAA AGACCAAAAG TGACTGGTGT CTCGTGTGAC AGATTGCTTC ATTTATGTTT 2100
CTACATAGTA AGGTGACTGC CAAATAATAT TTGAAGTCAT CTGTCTCTTT GTAAATTAAT 2160
TTATATGACC TATAAATTTA AAAATGTTTT TCAGTGAGTG CTTTAAACAA ACTTAAGCTT 2220
CTGCCCTGCC AAGGGAATTA ATGTTATCTT GTGAAAGGTG TTGCTGTTTG AATTGATGAG 2280
AAATGGAAGA TGAGAACTCC CTAAGAGTTC TCATAATAAA TCATCTCATC ACAAATCAAT 2340
ACGGTATACA GAGTTAAAGT GGAATGAGGT AAGAAGATAC AGCTACAGAA AATAGTTGCG 2400
TGATATGGAG AACAGTCATT GTAATTGGGT AGTTTGTGTA ATAAATATTT TTAATCTTTG 2460
CTTTTCAGAA ATTACCGAAT GTGTATAAAC AAATAAAGAA AAATAATTTA GCTGTGTTTT 2520
AGACAGCATT AGAATATATT GTTCAGCACA GTAAATATA TTGAAATTT GATAAGCCAA 2580
AAATGTGGTT TTGAATGAAT ATTTTGTGAA TCTTCTTAA AAGCTCAAAT TTGTAGACTT 2640
CTAAATAGAA TAAACACTTG CAGCAGAAAA AAAAAAATAA AAAAAAATAA AAAAAAATAA 2700
AAAAAATAA AAAAAAATAA AAAAAAATAA AAAAAAATAA AAAAAAATAA AAAAAAATAA 2760
AAAAAATAA AAAAAAATAA AAAAAAATAA AAAAAAATAA AAAAAAATAA AAAAAAATAA

```

SEQ ID NO:257 PBM1 Protein sequence:

PBM1 Protein sequence: CAB76901

60
65
70

```

MGDPSKQDIL TIFKRLRSLV TNKVCDFDCG KNPSWASITY GVFLCIDCSG SHRSLGVHLS 60
FIRSTELDSN WSWFQLRCMQ VGGNASASSF FHQHCSTND TNAKYNRAA QLYREKIKSL 120
ASQATRKHGT DLWLDSCVVP PLSPPKKEED FFASHVSPEV SDTAWASALA EPSSLTSRPV 180
ETTLNENEGG EQQGPSVEGL NVPTKATLEV SSIKKKKPNQ AKKGLGAKKG SLGAQKLANT 240
CFNIEIKQAA AADKMKEQED LAKVVSKEES IVSSSLRLAYK DLEIQMKKDE KMNISGKKNV 300
DSDRLGMGFG NCRSVISHSV TSDMQTIEQE SPIMAKPRKK YNDDSDSYF TSSSYFDEP 360
VELRSSFSF WDDSSDSYWK KETSKDTETV LKTTGYSDRP TARRKPDYEP VENTDEAQKK 420
FGNVKAISSD MYFGRQSQAD YETRARLERL SASSSISSAD LFEEPRKQPA GNYSLSSVLP 480
NAPDMAQFKQ GVRVAVGKLS VFANGVVTSI QDRYGS

```

SEQ ID NO:258 PBM4 DNA sequence

Nucleic Acid Accession#: D30891

Coding sequence: 1-4032 (underlined sequence corresponds to start and stop codon)

75
80

```

ATGGATACTG TCATGAAGCA GACACATGCT GACACACCTG TTGATCATTG TCTATCTGGC 60
ATAAGAAAGT GTAGCAGCAC CTTTAAGCTT AAAAGTGAAG TCAACAAGCA TGAACAGGCC 120
CTTGAAATGC AGAATCCAAA TTGAAACAAT AAAGAATGTT GTTTCACCTT TACGTTGAAT 180
GGAACTCCCA GAAATTTAGA CCGTAGTGTG TTTACAGCAT ATGGTAAACC CAGCGAGAGT 240
ATCTACTCAG CCTGAGTGC TAATGACTAT TTCAGTGAAA GGATAAAGAA TCAGTTTAAT 300
AAGAACATTA TTGTTTATGA AGAAAGACA ATAGATGGAC ATATAAATTT AGGAATGCCT 360
CTCAAGTGCC TGCTTATGTA TTCTCATTTT AAAATTACAT TTGGTCAAAG AAAGAGTAGC 420

```

AAAGAAGATG GACACATATT ACGCCAATGT GAAAAATCCAA ACATGGAATG CATTCTTTTT 480
 CATGTTGTGT CTATAGGAAG GACAAGAAAG AAGATTGTGA AGATCAACGA ACTTCATGAA 540
 AAAGGAAGTA AACTTTGTAT TTATGCCTTG AAGGGTGAGA CTATTGAAGG AGCCTTATGC 600
 AAGGATGGCC GTTTCCTGGT TGACATAGGT GAATTTGAAT GGAACTAAA GGAAGGTCAT 660
 AAGAAAATTT ATGGAAAAACA GTCCATGGTG GATGAAGTAT CTGGAAAAAGT CTTAGAAATG 720
 GACATTTCAA AAAAAAAGC ATTACAACAG AAAGATATCC ATAAAAAAT TAAACAGAAT 780
 GAAAGTGCCA CTGATGAAAT TAATACCAG AGTCTGATAC AGTCTAAGAA AAAAGTCCAC 840
 AAACCAAGA AAGATGGAGA GACCAAGAT GTAGAACACA GCAGAGAGCA AATTCTCCCA 900
 CCTCAGGATC TAAGCCATTA TATTAAAGAT AAAACTCGCC AGACAATTCC CAGGATTAGA 960
 AATTATTACT TTTGTAGTTT GCCCGAAAA TATAGGCAAA TAAACTCACA AGTTAGACGG 1020
 AGGCCGCATC TGGGTAGGCG GTATGCTATT AATCTGGATG TCCAAAAGGA GGCAATTAAT 1080
 CTCTTAAAGA ATTATCAAC GTTGAATGAA GCCATAATGC ATCAGTATCC GAATTTTAAA 1140
 GAGGAGGCAC ATGTGGTAAG AAAATATTTT CGGAAGAAG AAAAGAGAAT GAATCTTTCA 1200
 CCAGCTAAGC AATTCAACAT ATATAAAAG GACTTCGGAA AAATGACTGC AAATCTGT 1260
 TTAGTTGCAA CCTGCGAACA GCTTACATAT TATAGCAAGT CAGTTGGGTT CATGCAATGG 1320
 GACAATAATG GAAACACAGG TAATGCTACT TGCTTTGTCT TCAATGGTGG TTATATTTTC 1380
 ACCTGTGAC ATGTTGTACA TCTTATGGTG GGTAAAAACA CACATCCAAG TTTGTGGCCA 1440
 GATATAATTA GCAAAATGTC GAAGTAACC TTCATTATA CAGAGTTCTG CCCTACTCCT 1500
 GACAATTGGT TTTCCATTGA GCCATGGCTT AAAGTGTCCA ATGAAAAATCT AGATTATGCC 1560
 ATTTTAAAC TAAAGAAAA TGGAAATGCG TTTCTCCAG GACTATGGCG ACAGATTCT 1620
 CCTCAACCAT CTACTGGTTT GATTATTTA ATTGGTCATC CTGAAGGCCA GATCAAGAAA 1680
 ATAGATGGTT GTACTGTAT TCTCTAAAC GAACGATTGA AAAAATATCC AAACGATTGT 1740
 CAAGATGGGT TGGTAGATCT CTATGATACC ACCAGTAATG TATACTGTAT GTTTACCCAA 1800
 AGAAGTTCC TATCAGAGGT TTGGAACACA CACACGCTTA GTTATGATAC TTGTTTCTCT 1860
 GATGGGTCTC CAGGCTCCCC AGTGTTTAAT GCATCTGGCA AATTGGTTGC TTGTCATACC 1920
 TTTGGGCTTT TTTATCAACG AGGATTAAAT GTGCATGCC TTATGAAAT TGGTTATTCT 1980
 ATGGATTCTA TCTTTGTGA TATTAAAAAG ACAATGAGA GCTTGTATA ATCATTAAAT 2040
 GATGAGAAAC TTGAGACCTA CGATGAAGAG AAAGCCCGGC CCAGGCCAGC CTACCGGCGA 2100
 CTAGGATGCT TTCGTTTTCG CTCTCGCTT CCAATACTCG GGACTGGGGA AACCGGGAGA 2160
 ATAGAAGCAG GCAAGGACCG CCGTGGGCAC GGGGTCAAGT AGACAGGGTC CTGCTCGCGG 2220
 CGTCAAGGAG GCAGGCTGTG GGTGCCCA GCGCAGCCAA TCGGCTCCG AAGTAGCTGG 2280
 AGCTCTGGAG CCTTTGCTTC CTCAAATACG AGCGGGAAC GCGTTGAGCG CTGGATTCCA 2340
 GGCCGAGTGC TGGCGAGGCG CGCAGTCTCT AAAGAGCAAC AGAATAATTG CAGTACTTCT 2400
 CTAATGAGGA TGGAGTCTAG AGGAGACCCA AGAGCCACAA CTAATACCCA GGCTCAAAGA 2460
 TTCCATTAC CTAAGAAAAA TCCAGAAGAC CAGACCATGC CCAAAAATAG GACAATATAT 2520
 GTTACCTTGA AGGCTGTAC AAAAGAGATA GAACTCACC AAGGCCAAGA AATGCTTTGT 2580
 CGTGGCACAG AAGGAATCAA AGAGTACATA AACCTTGGAA TGCCCTCAG TTGTTTCCCT 2640
 GAAGGTGGCC AGGTGGTCAAT TACATTTTCC CAAAGTAAAA GTAAAGCAGAA GGAAGATAAC 2700
 CACATATTG GCAGGACGAA CAAAGCATCG ACTGAATGTG TCAAATTTA CATTATGCA 2760
 ATTGGAATTG GGAAGGTAA AAGAAGGATT GTTAAATGTG GGAAGCTTCA CAAAAAGGG 2820
 CGCAAATCT GTGTTTATGC TTCAAAGGA GAAACCATCA AGGATGCACT GTGCAAGGAT 2880
 GGCAAGATT TTTCTTTCT GGAGAATGAT GATTGGAAC TCATTGAAAA CAATGACACC 2940
 ATTTTAGAAA GCACCCAGCC AGTTGATGAA TTAGAAGGCA GATACTTTCA GGTGAGGTT 3000
 GAGAAAAAGAA TGGTCCCCG TGCAGCAGCT TCTCAGAATC CTGAGTCAGA GAAAAAGAAC 3060
 ACCTGTGTGT TGAGAGAAC AATCGTGGCT CAGTACCCCA GTTTGAAAA AGAAAGTGAA 3120
 AAAATCATTT AAAACTTCAA GAAAAAATG AAAGTAAAAA ATGGGGAAAC ATTATTGAA 3180
 TTGATAGAA CAACGTTTGG GAAAGTAA CAAAAATCTT CTTCGATTAA AGTAGTGAA 3240
 CTCTTTGAC GTCTCAGTGA CTCAGTTGG TACTTATTCT GGGACAGTGC AACTACGGGT 3300
 TACGCCACCT GCTTTGTTTT TAAAGGATTG TTCATTITAA CTGTGCGCA TGTAATAGAT 3360
 AGCATTTGG GAGACGGAAT AGAGCCAAGT AAGTGGGCAA CCATAATTGG TCAATGTGTA 3420
 AGGGTGACAT TTGGTTATGA AGAGCTAAAA GACAAGGAAA CAAACTACTT TTTTGTGAA 3480
 CCTTGGTTTG AGATACATAA TGAAGAGCTT GACTATGCTG TCCTGAACT GAAGGAAAAAT 3540
 GGACAACAAG TACCTATGGA ACTATATAAT GGAATTACT CTGTGCACT TAGTGGGTTG 3600
 ATACATATTA TTGGCCATCC ATATGGAGAA AAAAAGCAGA TTGATGCTTG TGCTGTGATC 3660
 CCTCAGGTC AGCGAGCAA GAAATGTGAG GAACTGTGCT AGTCTAAAAA AGCAGAAAGT 3720
 CCAGAGTATG TCCATATGTA TACTCAAAGA AGTTTCCAGA AAATAGTTCA CAACCTGAT 3780
 GTGATTACCT ATGACACTGA ATTTTCTTT GGGGCTTCCG GCTCCCTGT GTTTGATTCA 3840
 AAAGGTTTCA TGGTGGCCAT GATGCTGCT GGCCTTGCT ATACTACCA AAATGAGACT 3900
 CGTAGTATCA TTGAGTTTGG CTCTACCATG GAATCCATCC TCCTGTATAT TAAGCAAAGA 3960
 CATAAACCAT GGTATGAAGA AGTATTTGTA AATCAGCAGG ATGTAGAAAT GATGAGTGAT 4020
 GAGGACTTGT GAGAATTCAG TCTACTGGAT TTAAGGGAAT GGCTTATGGA GTTGTTATTT 4080
 CGTAGGCATT GAAAATGGTT TTCTAAACTC CAAAATGGTC ATCTTATCAA TAATAATAAT 4140
 ATTGACCAT TCTATCTGC CAGGCATTTT TCTAAGCACA TGAAGAAATT AGTCCTAACA 4200
 ACACTATGAG ATGGACTATA ACTTGCCCAA ATTTTTTTTT TTTTGTAGAC TGAGTCTCAC 4260
 TCTGTGCGCT GGGCTGGAGT ACAGTGGTGC GATCTCAGCT CACTGCACT TCCACCTCCC 4320
 AGGTTCAAAG GATTCTTATG CTCAGTCTC CTGAGCAGCT GGGATTACAG GCAAACGCCA 4380
 CCACACCCAG CTAATTTTTT TTTTTTTTTT TGTATTTTGA GTAGAGACAG GGTTTCACCA 4440
 TGTGTTGTCAG GCGGGTCTCG AACTCCTGAC CTCGTGATCC ACCTGCCTCG GCCTTCCAAA 4500
 GTGCTGGGAT TACAAGTTTG AGCCACTGCA CCTGGCTAAC TTGCCCTATT TTAAGTCAA 4560
 GCAATGGGAA GAATAACAAG ATTATATAGT AATCAGTTTC ATGACACTAA AAGTCATATA 4620
 GTCATAGGT TTTTCTCT TCCATATCT TGCCTAAAT CATTGTCTAC AGTGCAGGA 4680
 CCAAACTTGT TTCATCTCAT GATTCCCTAC ATCTGACATA AGGAAAGTAA GTGCTCAGAA 4740
 AAATGTGAG GTCAATAAGT TGCAAAAGT GGGGCTGCAA TTAATGCTAA CATAAGAGCT 4800
 AAATGCTTGA TTAGAAATGA TCTCAAAACC TTTTGAAT TCCAAATCT TCATATTACT 4860
 GAAACTGTG GAATATATGG GTCCTGAAAT TCAGAAGATG ATAGTCACTC TTCCCATATT 4920
 TATAGGCTAT TAAGGCAAGG GATATCTTAA ACATCATATT ACTTTATTTA GATTCTACT 4980
 ACTCCAATTA TTAATGTTAT GTATTCTCA TTGTTTACT TCTCATGGT ATTATGAAGA 5040
 CTATATAGAT GATTCAACCA AGCCTGCAAA TCTCCCTCT GTGGAATTC ACTGGACCCA 5100
 ATCTGTTTTT CATTTCATT GCAATACTAC TAAAGCCATA CAATATCAAG CACCTCCCT 5160

CTAGGTCCAG GGAATATCAC AGAAGAAGCA GGCATGTAAG ATTTTAAGGA CTGGTTTCGA 5220
 GGGGTCGAGT GTAGGAAAAC AGCCTGTTGC ATTGTAAGAG TGATGTCACC TTGAAGAGCA 5280
 GCTGGCATGA TGACTGCTGT TTGACTCCTG CATACCAAGA TATTCTGCAG CAATGTCTTT 5340
 AAACAGTGCC GGTAGTACAG ATAACCCCTC ATAAAGATGC TTATCTAACC TCCCCAGTGT 5400
 TCAGGTGTTT CACAAGAAAG TCTGAGATAT GACTAGCTAC ACGTTTTGCC AAAAATGCTT 5460
 GTTATATAAA GGGTACTTTT GGGAGGGTGA GTGCCGCCAT TTAGTGGCTG CTAGAAACAT 5520
 TGCTTCTGTT TGTAAGTTCC TATTAAATGT TCTTCTGAG AAAAAAAAAA A

SEQ ID NO:259 PBM4 Protein sequence:
 PBM4 Protein sequence: BAB67788

MDTVMKQTHA DTPVDHCLSG IRKCSSTFKL KSEVNKHETA LEMQNPNLNN KECCTFTTLN 60
 GNSRKLD RSV FTAYGKPS ES IYSALSANDY FSERIKNQFN KNIIVYEEKT IDGHINLGMP 120
 LKCLPSDSHF KITFGQRKSS KEDGHILRQC ENPNMECILF HVVAIGRTRK KIVKINELHE 180
 KGSKLCIYAL KGETIEGALC KDGRFRSDIG EFEWKLKEGH KKIYGKQSMV DEVSGKVLEM 240
 DISKKKALQQ KDIIKKIKQN ESATDEINHQ SLIQSKKKVH KPKKDGETKD VEHSREQILP 300
 PQDLSHYIKD KTRQIPRIR NYFFCSLPRK YRQINSQVRR RPHLGRRYAI NLDVQKEAIN 360
 LLKNYQITLNE AIMHQYPNFK BEAQWVRKYF REEQKRMNLS PAKQFNYYKK DFGKMTANSV 420
 SVATCEQLTY YKSVGFQWQW DNNNGTGNAT CFVFNGGYIF TCRHVHLMV GKNTHPSLWP 480
 DIISKCAKVT FTYTEFCPTP DNWFSIEPWL KVSNNENLDYA ILKLKENGNA FPPGLWRQIS 540
 PQPSTGLIYL IGHEPGQIKK IDGCTVIPLN ERLKKYPNDC QDGLVDLYDT TSNVYCMFTQ 600
 RSFLSEVWNT HTLSYDTCFS DGSSGSPVFN ASGKLVALHT FGLFYQRGFN VHALIEFGYS 660
 MDSILCDIKK TNESLYKSLN DEKLETYDEE KARPRPAYRR LGCFRFRSRF PILGTGETGR 720
 IEAGKDRRGH GVSETGSCSR RQGGALWVSP AQPIGRSSW SSGAFASNT SGNCVERWIP 780
 GRVLARRAYS KEQQNNCSTS LMRMESRGDP RATTNTQAQR FHSPKKNPED QTMPQNRTTY 840
 VILKAVRKEI ETHQGQEMLV RGTEGIEKEYI NLGMPLSCFP EGGQVVITFS QSKSKQKEDN 900
 HIFGRQDKAS TECVKFYIHA IGIGCKRRI VKCGKLHKKG RKLVCYAFKG ETIKDALCKD 960
 GRFLSFLEND DWKLIENNDT ILESTQPVDE LEGRYFQVEV EKRMVPSAAA SQNPSEKRN 1020
 TCVLREQIVA QYPSLKRESE KIENFKKKM KVKNGETLFE LHRTTFGKVT KNSSSIKVVK 1080
 LLVRLSDSVG YLFWDSATTG YATCFVFKGL FILTCRHVID SIVGDGIEPS KWATHIGQCV 1140
 RVTFGYEELK DKETNYFFVE PWFEIHNEEL DYAVLKLKEN GQQVPMELYN GITPVPLSGL 1200
 IHIHGPYGE KKQIDACAVI PQGQRAKKCQ ERVQSKAES PEYVHMYTOR SFQKIVHNPD 1260
 VITYDTEFFF GASGSPVDS KSLVAMHAA GFAYTYQNET RSIEFGSTM ESILLDIKQR 1320
 HKPWYEEVFV NQQDVEMMSD EDL

SEQ ID NO:260 PBQ1 DNA sequence
 Nucleic Acid Accession#: NM_015642
 Coding sequence: 489-2489 (underlined sequence corresponds to start and stop codon)

1	11	21	31	41	51	
ACATTTCAAA	AAAAATACAT	AGACTGATGT	TTGACAGTTG	TGCAGCATAA	GCCTACAGGG	60
TACGAAGAA	GAACCTCTGAG	AATGTTTGGA	GAATGTTTCA	TCATTACTAA	CAGGATATTC	120
CTCATGACAT	TGCTGTCTGA	TCTTTGACCA	TCAGTCTGTG	ACCTGCCCTT	TCTCTTTACA	180
TGCAGCCGCT	CTCTGTCTCC	TGCCCAATG	AACATCTGCA	CTAGGCCCAA	GCCTTGGAGT	240
AATTTACCTG	AAGAGTGACA	CCATTGATTT	TGAAACTACT	GAAGAAACCC	AAGACAGCTG	300
AAAACCAAGAA	GGCATCTGAG	GAGAATGAGA	TTACTCAGCC	GGGTGGATCC	AGCGCCAAGC	360
CGGGCCTTCC	CTGCCTGAAC	TTTGAAGCTG	TTTGTCTCC	AGACCCAGCC	CTCATCCACT	420
CAACACATTC	ACTGACAAAC	TCTCAGCTC	ACACCGGTC	ATCTGATTGT	GACATCAGTT	480
GCAAGGGGAT	GACCGAGCGC	ATTCACAGCA	TCAACCTTCA	CAACTTCAGC	AATTCCTGTC	540
TCGAGACCTT	CAACGAGCAG	CGCAACCGTG	GCCACTTCTG	TGACGTAACG	GTGCGCATCC	600
ACGGGAGCAT	GCTGCGCGCA	CACCGCTGCG	TGCTGGCAGC	CGGCAGCCCC	TTCTTCCAGG	660
ACAAACTGCT	GCTTGGCTAC	AGCGACATCG	AGATCCCCTC	GGTGGTGTCA	GTGCAGTCAG	720
TGCAAAAGCT	CATTGACTTC	ATGTACAGCG	CGGTGCTACG	GGTCTCGCAG	TCGGAAGCTC	780
TGCAGATCCT	CACGGCCGCC	AGCATCCTGC	AGATCAAAAC	AGTCATCGAC	GAGTGCACGC	840
GCATCGTGTC	ACAGAACGTG	GGCGATGTGT	TCCCGGGGAT	CCAGGACTCG	GGCCAGGACA	900
CGCCGCGGGG	CACCTCCGAG	TCAGGCACGT	CAGGCCAGAG	CAGCGACACG	GAGTCGGGCT	960
ACCTGCAGAG	CCACCCACAG	CACAGCGTGG	ACAGGATCTA	CTCGGCACCT	TACGCGTGCT	1020
CCATGCAGAA	TGGCAGCGGC	GAGCGCTCTT	TTTACAGCGG	CGCAGTGGTC	AGCCACCACG	1080
AGACTGCGCT	CGGCTTGCCC	CGCGACCACC	ACATGGAAGA	CCCCAGCTGG	ATCACACGCA	1140
TCCATGAGCG	CTCGCAGCAG	ATGGAGCGCT	ACCTGTCCAC	CACCCCGCAG	ACCACGCACT	1200
GCCGCAAGCA	GCCCCGGCCT	GTGCGCATCC	AGACCCTAGT	GGGCAACATC	CACATCAAGC	1260
AGGAGATGGA	GGACGATTAC	GACTACTACG	GGCAGCAAAG	GGTGCAGATC	CTGGAACGCA	1320
ACGAATCCGA	GGAGTGACAG	GAAGACACAG	ACCAGGCCGA	GGGCACCGAG	AGTGAGCCCC	1380
AAGGTGAAAG	CTTCGACTCG	GGCGTCAGCT	CCTCCATAGG	CACCGAGCCT	GACTCGGTGG	1440
AGCAGCAGTT	TGGGCTTGGG	GCGGCGCGGG	ACAGCCAGGC	TGAACCCACC	CAACCCGAGC	1500
AGGCTGCAGA	AGCCCCCGCT	GAGGGTGGTC	CGCAGACAAA	CCAGCTAGAA	ACAGGTGCTT	1560
CCTCTCCGGA	GAGAAGCAAT	GAAGTGGAGA	TGGACAGCAC	TGTTATCACT	GTGAGCAACA	1620
GCTCCGACAA	GAGCGTCTTA	CAACAGCCCT	CGGTCAACAC	GTCCATCGGG	CAGCCATTGC	1680
CAAGTACCCA	GCTTACTTTC	CGCCAGACAG	AAACCCCTAC	CAGCAACCTG	AGGATGCCTC	1740
TGACCTTGAC	CAGCAACACG	CAGGTCAATTG	GCACAGCTGG	CAACACCTAC	CTGCCAGCCC	1800
TCTTCACTAC	CCAGCCCGCG	GGCAGTGGCC	CCAGCCCTTT	CCTCTTCAGC	CTGCCACAGC	1860
CCCTGGCAGG	CCAGCAGACC	CAGTTTGTGA	CAGTGTCCCA	GCCCCGTCTG	TCGACCTTTA	1920
CTGCACAGCT	GCCAGCGCCA	CAGCCCCCTG	CCTCATCCCG	AGGCCACAGC	ACAGCCAGTG	1980
GGCAAGGCGA	AAAAAAGCCT	TATGAGTGCA	CTCTCTGCAA	CAAGACTTTC	ACCGCCAAAC	2040
AGAACTACGT	CAAGCACATG	TTCTGTACACA	CAGGTGAGAA	GCCCCACCAA	TGCAGCATCT	2100
GTTGGCGCTC	CTTCTCCTTA	AAGGATTACC	TTATCAAGCA	CATGGTGACA	CACACAGGAG	2160

5 TGAGGGCATA CCAGTGTAGT ATCTGCAACA AGCGCTTCAC CCAGAAGAGC TCCCTCAACG 2220
 TGCACATGCG CCTCCACCGG GGAGAGAAGT CCTACGAGTG CTACATCTGC AAAAAGAAGT 2280
 TCTCTCACAA GACCCCTCCTG GAGCGACACG TGGCCCTGCA CAGTGCCAGC AATGGGACCC 2340
 CCCCTGCAGG CACACCCCA GGTGCCCGCG CTGGCCCCC AGGCGTGGTG GCCTGCACGG 2400
 AGGGGACCAC TTACGTCTGC TCCGTCTGCC CAGCAAAGTT TGACCAAATC GAGCAGTTCA 2460
 ACGACCACAT GAGGATGCAT GTGTCTGACG GATAAGTAGT ATCTTTCTCT CTTTCTTATG 2520
 AACAAAACAA AACAAACAACA AAAACAACA AAACAAAAA GCTATGGCAC TAGAATTTAA 2580
 GAAATGTTTT GGTTCATTT TTAATTTCTG TTTTGTGTTT TGTTTCGTTT CATTTGTGAC 2640
 10 TACATGAAGA ACTGTTTTTT GCCTGCTGGT ACATTACATT TCCGGAGGCT TGGGTGAATA 2700
 ATAGTTTTC CAGTCTCCCT CGGATGGTGG CCTTAAGGCC TGGTAGTGCT TCAAGAGGTC 2760
 CACTGGTGG ATCTCTAGCT ACTGGCCTCT AAATACAACC CTTCCTTACA AAAAAAAA 2820
 AAAAAAAA

15 SEQ ID NO:261 PBQ1 Protein sequence:
 PBQ1 Protein sequence: NP_056457

20 MTERIHSINL HNFNSVLET LNEQRNRGHF CDVTVRIHGS MLRAHRCVLA AGSPFFQDKL 60
 LLGYSIDIEP SVVSVQSVQK LIDFMYSGVL RVSQSEALQI LTAASILQIK TVIDECTRIV 120
 SQNVGDVFPQ IQDSGQDTPR GTPESGTSQ SSDESGLYLQ SHPQHSVDRI YSALYACSMQ 180
 NGSGERSFYS GAVVSHHETA LGLPRDHME DFSWITRIHE RSQOMERYLS TTPETTHCRK 240
 QPRPVRIQL VGNIIHKQEM EDDYDYYGQ RVOILERNES EECTEDTDQA EGTESEPKGE 300
 SFDSGVSSSI GTEPDSVEQQ FGPGAARDSQ AEPTQPEQAA EAPAEGGPQT NQLETGASSP 360
 25 ERSNEVEMDS TVITVNSSD KSVLQQPSVN TSIGQLPST QLYLRQITEL TSNLRMLPTL 420
 TSNTQVIGTA GNTYLPALFT TQAPGSGPKP FLFSLPQLA GQQTQFVTVS QPGLSTFIAQ 480
 LPAPQLASS AGHSTASGQG EKKPYECTLC NKTFTAKQNY VKHMFVHTGE KPHQCSICWR 540
 SFSLKDYLIK HMTVHTGVRA YQCSICNKR TQKSSLNVMH RLHRGEKSYE CYICKKKFSH 600
 KTLERHVAL HSASNGTPPA GTPPGARAGP PGVVACTEGT TVCVSCPAPK FDQIEQFNH 660
 30 MRMHVSDG

35 SEQ ID NO: 262 PBQ6 DNA sequence
 Nucleic Acid Accession#: AI654187
 Coding sequence: 1-912 (underlined sequence corresponds to start and stop codon)

40 1 11 21 31 41 51
 | | | | | |
 ATGGTGGAAG AGGAAACAGG CATATCTTAC ATGGTGCCAG ACAAGGGACA CCCTTCTACA 60
 AACTCTACCA CTCTGCGGCC GTCGTTTCGA CCATATAAAA ACGACCTATG CGAACTGCGT 120
 CGGAAACTC CCTCAGCATG TAAACGAAG ATCAGGAGCA GATTTGAAGA ATTACAAAGT 180
 GAATTGGTGC CAGTCAGCAT GTCAGAGACA GACCACATAG CCTCTACTTC CTCTGATAAA 240
 AATGTTGGGA AAACACCTGA ATTAAAGGAA GACTCATGCA ACTTGTCTTC TGGCAATGAA 300
 AGCAGCAAT TAGAAAATGA GTCCAAACTA TTGTCAATTA ACACGTATAA AACTTTATGT 360
 45 CAACCTAATG AGCATAATAA TCGAATTGAA GCCCAGGAAA ATTATATTCC AGATCATGGT 420
 GGAGGTGAGG ATTCTTGTGC CAAACAGAC ACAGGCTCAG AAAATCTCTGA ACAAATAGCT 480
 AATTTTCCTA GTGGAATTT TGCTAAACAT ATTTCAAAAA CAATGAAAC AGAACAGAAA 540
 GTAACACAAA TATTGTGGA ATTAAAGTCA TCTACATTT CAGAATCAGC TAATGAAAAG 600
 ACTTATTTCG AAAGCCCTTA TGATACAGAC TGACCAAGA AATTATTTTC AAAAATAAAG 660
 50 AGCGTTTCAG CATCAGAGGA TTTGTTGGA GAAATAGAAT CTGAGCTCTT ATCTACGGAG 720
 TTTGCAGAAC ATCAGTACC AAATGGAATG AATAAGGAG AACATGCATT AGTTCTGTTT 780
 GAAAGTGTG TGCAAGATAA ATATTGTCAG CAGGAACATA TCATAAAAA GGCCAGACTT 840
 GGTCTCTGTT ATTGTCATC AAGAACCTCA ATTGACACGT TAATCCGTT TATCCCAAT 900
 TTATATAGAT AA

55 SEQ ID NO:263 PBQ6 Protein sequence:
 Protein Accession #: NP_060170

60 MEPKEATGKE NMVTKKKNLA FLRSRLYMLE RRTDITVVES SVSGDHSGTL RRSQSDRTEY 60
 NQKLQEKMTQ QGECVSAETL TPREEHHMKR MMAKREKIHK ELIQTEKDYL NDLELCVREV 120
 VQPLRNKKT DRLDVSLSFN IESVHQISAK LLSLLEAATT DVEPAMQVIG EVFLQIKGPL 180
 EDIYKIYCYH HDEAHSILES YEKEELKEH LSHCIQSLK

65 SEQ ID NO:264 PBQ7 DNA sequence
 Nucleic Acid Accession#: NM_014323
 Coding sequence: 662-2725 (underlined sequence corresponds to start and stop codon)

70 1 11 21 31 41 51
 | | | | | |
 GGGCTACTC TGCCGCCGCC GCCGCCGCC CGCTCCAGCC GCGCGCCCG CCGCCACCGC 60
 CCTCCAGGT CCGGGACCGG GCCCGGCCCA CCGCCCCCGT GCGCGCCCG CCGCCGCCGC 120
 CTTCGCCCTT GCCTTTGTTT TCTTCCGCTC CGCGCGCCCG GCGCGCGCT CCGCTTTGCA 180
 75 GGGACGCGAG CGCGCGCCC CAGCGGGCCC GGGAAAGGCC GCGCGCGCG CCGCGCGCTG 240
 CGCGCGGAC CCTTCTCTCT CCTCCCCGCG TGCGCGTGCC CTCTTGGCT GCGCGCGCG 300
 GCCGCTTGC GGGCGGGAGG GGAGGTGGCA GCGCGCTTTG CAGGAGGGGC GCACCTCTTC 360
 GCTCGCGCAC CCCCCCGGAA GGTAGACCGG GAAGGGGAGG CCGCGCGCGG GAGAGGAGAG 420
 AGTGCGCGCG AGTCCAGGTA GGGCGGGGGT TGGCTATGTG GGGGGTGGTG CACCCGCGAG 480
 80 TCTAGACAGT CTGATCCGGG CTGGGGCGGT GTACACTCGG CGCACCTCGG AGACTACAGA 540
 GCCTCGGGCC GGCACGTGTG GGGAGTGTGG ACACGTCTGC TCGCGCCCGC TTCTCGCTGC 600

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

TGAGGGGAAG GAGGGGGCG GGCAGGTGCA GCGGCCGGGC TAGTGGGAGG GGGCGGCGGC 660
 CATGGAGCGG GTGAACGACG CTTCGTGCGG CCCGTCTGGC TGCTACACAT ACCAGGTGAG 720
 CAGACACAGC ACGGAGATGC TGCACAACCT GAACCAGCAG CGCAAAAACG GCGGGCGCTT 780
 CTGCGACGTG CTCTTGCGGG TAGGCGACGA GAGCTTCCCA GCGCACCGCG CCGTGTCTGGC 840
 CGCCTGCAGC GAGTACTTTG AGTCGGTGTG CAGCGCCAG TTGGGCGAGC GCGGAGCTGC 900
 GGACGGGGGT GCGGCTGATG TAGGGGGCGC GACGGCAGCA CCAGGCGGCG GGGCGGGGGG 960
 CAGCGGGGAG CTGGAGATGC AACTATCAG CTCCAAGGTA TTGGGGGACA TTCTGGACTT 1020
 CGCCTACACT TCCCGCATCG TGGTGCCTT GGAGAGCTTT CCCGAACCTCA TGACGGCCGC 1080
 CAAGTTCTCTG CTGATGAGGT CGGTTATCGA GATCTGCCAG GAAGTCATCA AACAGTCCAA 1140
 CGTACAGATC CTGTTACCCC CTGCCGCGC CGATATAATG CTCTTTCGCC CCCCTGGGAC 1200
 CTCGGACTTG GGGTTCCTT TGGACATGAC CAACGGGGCA GCCTTGGCAG CCAACAGCAA 1260
 TGGCATCGCC GGCAGCATGC AGCCAGAGGA GGAGGCGACT CCGGCGGCTG GTGCAGCCAT 1320
 TGCAGGGCAA GCCTCTTTGC CTGTGTATAC TGGGGTGGAC CGCTTGCCCA TGGTGGCTGG 1380
 ACCCTATACC CCCCAACTGC TGACTTCCCC ATTCGCCAGT GTGGCATCCA GTGCCCTCC 1440
 CTGACTTGC AAGCGAGGCC GGGGCCGCCC AAGGAAGGCC AACCTGCTGG ACTCAATGTT 1500
 TGGGTCCCCA GGGGGCTGA GGGAGCGAG CATCTTCCA TGCGGTCTAT GTGGTAAGGT 1560
 GTTCACTGAT GCCAACCGGC TCCGGCAGCA CGAGGCCAG CACGGTGTCA CCAGCTTCCA 1620
 GCTGGGTAC ATCGACTTC CTCTCCGAG GCTGGGTGAG AATGGGTAC CCATCTCTGA 1680
 AGACCCGAC GGGCCCGAA AGAGGAGCC GACCAGGAAG CAGGTGGCTT GTGAGATCTG 1740
 CGGCAAGATC TTCCGTGATG TGTATCATCT TAACGGCAC AAGCTGTCCC ACTCTGGGGA 1800
 GAAGCCCTAC TCCTGCCCTG TGTGTGGGT GCGGTTCAG AGAAAAGACC GCATGTCTTA 1860
 CCATGTGCGG TCCCATGATG GGTCCGTGGG CAAGCCTTAC ATCTGCCAGA GCTGTGGGAA 1920
 AGGCTTCTCC AGGCTGATC ACTTGAACGG ACATATCAAG CAGGTGCACA CTCTGAGCG 1980
 GCCTCAACAG TGTCAACCT GCAATGCTT TTTGCCACC CGAGACCGTC TCGCTCCCA 2040
 CCTGGCTGT CATGAAGACA AGGTGCCCTG CCAGGTGTGT GGGAGTACT TCGGGCAGC 2100
 ATACATGGCA GACCACCTGA AGAAGCACAG CGAGGGGCCC AGCAACTTCT GCAGTATCTG 2160
 TAACCGAGGT TTCTCCCTG CTCTCACTT AAAGTCCAT GTTAAACCC ACCACGGTGT 2220
 TCCCTTCCC CAGGTCTCCA GGCACAGGA GCCATCTCTG AATGGGGGAG CAGCGTTCCA 2280
 CTGCGCCAGG ACCTATGGCA ACAAGAAGG CCAGAAATGC TCACATCAGG ATCCGATTGA 2340
 GAGCTCTGAC TCCTATGGTG ACCTCTCAGA TGCCAGCGAC CTGAAGACGC CAGAGAAGCA 2400
 GAGTGCCAAAT GGTCTTTCT CTGCGACAT GGCAGTCCCC AAAAAACAAA TGGAGTCTGA 2460
 TGGGGAGAAG AAGTACCAT GCCTGAATG TGGGAGCTT TTCCGCTCTA AGTCTACTT 2520
 GAACAAACAC ATCCAGAAGG TGCATGTCCG GGCTCTCGGG GGGCCCTGG GGGACCTGGG 2580
 CCCTGCCCTT GGCTCACCTT TCTCTCTCA GCAGAACATG TCTCTCTCG AGTCTTTGG 2640
 GTTTCAGATT GTTCAGTGG CATTTGCGTC ATCTTAGTA GATCCTGAGG TTGACCAGCA 2700
 GCCCATGGG CCTGAAGGGA AATGAGGAG CTGCTGTGTC CCCACGGAAA CAACCATCTG 2760
 GGGACTGCTG GGAATGCTG TGAATGCGGA GGAAGTGAT GTTTGGGTTT TGTAGCTGAG 2820
 AGATTTTAT TCATTTTAA TGCCCCCA ACCCACTCC AACTCTTCT CCACCACCA 2880
 TTCTCCCAAT GGTCTTTAGA AATAGATTT CATCTGATAT TCTGCAGAA TATCAATGAG 2940
 ACTTGGTATG GGCAGGGGC AGAAAACT ACATAGGCCCT CCAAGGCAA ACCAGTCCA 3000
 GTTCTTTAA TGGGAAGAAG CTGGAATTC TGGTGTCTAA TTCTTAGTGA CCCCAATCT 3060
 ATACCCAAAT CTATGATATT CTGGACCTC AGTGATTTG GTCCCTCCC ACTTCTTAG 3120
 TTGCTCATCC TCCTTCCCA TATCCTTCAA AAGAACCACA CTAGGCTCT CACCTACTTA 3180
 TACAATGCGG ATGCCAACT GTTTTAAGG AAGCCAGAAG CATCCCATGG ACCATGGGGT 3240
 GAGTGTCTC CAAGAGCCC CTGAGCTCAG CCCTCTGCTT GAGGGCTCC AGACCTTCT 3300
 GAGCCCTGCT TGGAGGCGAG CATTTCACT GCTAGGACAA GCTCAGCTGT TGAGGACACC 3360
 CCCACCCCAA ATTTCAGTTC TTACGTGATT TTAACCATTC AACATGCTGT TGGGTTTAA 3420
 TTCTTAATT ATPATTTAT TGTATTAT TTTTAGGAC CAGTTGTAGT GAATTGCTAC 3480
 TGAAGCTAT CCAGGTGAT ACAGACTCT TGTAAACCG CAGTCACACA TTAGGGTTAG 3540
 TATTAACCT TGTTAGATG TACCATAATT AACTTGGCTA GTTGATTGTT TGAAGTCTAT 3600
 GGAAGAAATA GTTTATGCA AAATTTTAA AATGCCAGT CTGGTCAGGG AAGTAGGGG 3660
 TTTCATGCT GTTGGGAACC AGGAAGGTGG GACAGCCGC AGGTAGGGAC ATTGTGTACC 3720
 TCAGTTGTGT CACATGTGAG CAAGCCAGG TTGACCTTGT GATGTGAAAT GATCTGATCA 3780
 GACTGTATTA AAAATGTTAG TACATTACT TA

SEQ ID NO:265 PB7 Protein sequence:

Protein Accession #: NP_114439

60
 65
 70

MERVNDASCG PSGCYTYQVS RHSTEMLHNL NQQRKNGGRF CDVLLRVGDE SFPRAHRAVLA 60
 ACSEYFESVF SAQLGDGGAA DGGPADVGGA TAAPGGGAGG SRELEMHTIS SKVFGDILDF 120
 AYTSRIVVRL ESFPELMTAA KFLMRSVIE ICQEVIKQSN VQLVPPARA DIMLFRPPGT 180
 SDLGFPLDMT NGAALAANSN GIAGSMQPEE EAARAAGAAI AGQASLPVLP GVDRLPMVAG 240
 PLSPQLLTSP FPSVASSAPP LTGKRGRGRP RKANLLDSMF GSPGGLREAG ILPCGLCGKV 300
 FTDANRLRQH EAQHGVTSLQ LGYIDLPPPR LGENGLPISE DPDGPRKRSR TRKQVACEIC 360
 GKIFRDVYHL NRHKLSHSGE KPYSCPVCGL RFKRKDRMSY HVRSHDGSVG KPYICQSCGK 420
 GFSRPDHLNG HIKQVHTSER PHKCQTCNAS FATRDLRLSH LACHEDKVPC QVCGKYLRAA 480
 YMADHLKKHS EGPSNFCIS NREGQKCSHQ DPESSDSYG DLSHSDSLKT PEKQSANGSF 540
 SCDMAVPKNK MESDGEKKYP CPECGSFFRS KSYLNKHQK LVHVRALGGPL GDLGALGSP 600
 FSPQQNMSLL ESFGQIVQS AFASSLVDE VDQQPMGPEG K

SEQ ID NO:266 PB9 DNA sequence

Nucleic Acid Accession#: NM_012429

Coding sequence: 174-1385 (underlined sequence corresponds to start and stop codon)

75
 80

1 11 21 31 41 51
 | | | | |
 CCCTACTCCG CCTCTCGGGA TCCTTTAAGA GGCGGGGCTT GGCTGCCAGC TCCGCGGCCC 60
 GGGCAAAAGG CTGGGACTTT ACTCCGGGTG GCGGCGAGGA CGAGTCTGTG CTCCATCAGC 120

5	TGCCGCACCC	GCCGCCCTCCC	GCCCCCAAAC	CCCATCCCCG	CGGTTGAGCC	ACGATGAGCG	180
	GCAGAGTCGG	CGATCTGAGC	CCCAGGCAGA	AGGAGGCATT	GGCCAAGTTT	CGGGAGAAATG	240
	TCCAGGATGT	GCTGCCCGGCC	CTGCCGAATC	CAGATGACTA	TTTCTCCTTG	CGTTGGCTCC	300
	GAGCCAGAAG	CTTCGACCTG	CAGAAGTCGG	AGGCCATGCT	CCGGAAGCAT	GTGGAGTTCC	360
	GAAAGCAAAA	GGACATTGAC	AAACATCATTA	GCTGGCAGCC	TCCAGAGGTG	ATCCAACAGT	420
	ATCTGTCAAG	GGGTATGTGT	GGCTATGACC	TGGATGGCTG	CCCAGTCTGG	TACGACATAA	480
	TTGGACCTCT	GGATGCCAAG	GGTCTGTGT	TCTCAGCCTC	CAAACAGGAC	CTGCTGAGGA	540
	CCAAGATGCG	GGAGTGTGAG	CTGCTTCTGC	AAGAGTGTGC	CCACCAGACC	ACAAAGTTGG	600
10	GGAGGAAGGT	GGAGACCATC	ACCATAATTT	ATGACTGCGA	GGGGCTTGGC	CTCAAGCATC	660
	TCTGGAAGCC	TGCTGTGGAG	GCCTATGGAG	AGTTTCTCTG	CATGTTTGAG	GAAAAATTATC	720
	CCGAAACACT	GAAGCGTCTT	TTTGTGTGTA	AAGCCCCCAA	ACTGTTTCCT	GTGGCCTATA	780
	ACCTCATCAA	ACCCTTCTCTG	AGTGAGGACA	CTCGTAAGAA	GATCATGGTC	CTGGGAGCAA	840
	ATTGGAAGGA	GGTTTTACTG	AAACATATCA	GCCTTGACCA	GGTGCTGTG	GAGTATGGGG	900
15	GCACCATGAC	TGACCCCTGAT	GGAAACCCCA	AGTGCAAATC	CAAGATCAAC	TACGGGGGTG	960
	ACATCCCCAG	GAAGTATTAT	GTGCGAGACC	AGGTGAAACA	GCAGTATGAA	CACAGCGTGC	1020
	AGATTTCCCG	TGGCTCCTCC	CACCAAGTGG	AGTATGAGAT	CCTCTTCCCT	GGCTGTGTCC	1080
	TCAGGTGGCA	GTTTATGTCA	GATGGAGCGG	ATGTTGGTTT	TGGGATTTTC	CTGAAGACCA	1140
	AGATGGGAGA	GAGGCAGCGG	GCAGGGGAGA	TGACAGAGGT	GCTGCCCAAC	CAGAGGTACA	1200
20	ACTCCACCT	GGTCCCTGAA	GATGGGACCC	TCACCTGCAG	TGATCCTGGC	ATCTATGTCC	1260
	TGCGGTTTGA	CAACACCTAC	AGCTTCATTC	ATGCCAAGAA	GGTCAATTTC	ACTGTGGAGG	1320
	TCCTGCTTCC	AGACAAAGCC	TCAGAAGAGA	AGATGAAACA	GCTGGGGGCA	GGCACCCTCA	1380
	AATAACACCT	TCTCCTATAG	CAGGCCCTGGC	CCCCTCAGTG	TCTCCCTGTC	AATTTCTACC	1440
	CCTTGTAGCA	GTCAATTTTCG	CACAACCCCTG	AAGCCCCAAG	AAACTGGGCT	GGAGGACAGA	1500
25	CCTCAGGAGC	TTTCATTTTCA	GTTAGGCAGA	GGAAAGAGCGA	CTGCAGTGGG	TCTCCGTGTC	1560
	TATCAAATAC	CTAAGGAGTC	CCCAGGAGCT	GGCTGGCCAT	CGTGATAGGA	TCTGTCTGTC	1620
	CTGTAAACTG	TGCCAACTTC	ACCTGTCCAG	GGACAGCGAA	GCTGGGGGTG	GCGGGGGGCA	1680
	TGTACACAG	GGTGGCAGCA	GGGAAAAAAA	TTAGAAAAAGG	GTGAAAGATT	GGGACTTAAC	1740
	ACTTCAGGGA	AGTCAGCTGC	CGGGGAGAAA	CTTGCTCCTA	AATGAACACA	TAAGTTTAGA	1800
30	TCGCAATGAG	GAGTAGCAGG	GTAGCTGGTT	GCTAGAGTTA	CGGTGGGGAT	CAGAAACTCT	1860
	TCCAAACATT	TTAGCACTGA	GGCTGGGGTA	GCTTTTGGCT	TTTCCCAGGT	CTCAGGAGGT	1920
	GGCCTGAGTC	AGCACACATC	TTCCCACTCG	GTAGACAGGC	TGGCCTCTCC	CTCACTTTGA	1980
	GACTTTGGCA	ACTCCTGGGC	CACACGGCCT	GCTCTTTTGA	TTACTAATGA	TTGTCAGTGA	2040
	CTCAGAGCTT	CCTGGGACTT	CGGGTACCCA	CCCCTCTTTC	TCCATGCAAA	CAAAGCGCCA	2100
35	GGGAAATGAC	CCACAGGGAT	CGCAGCTGCA	GGGAGGGCCA	GGGAGGTGG	GGGTGGGAGT	2160
	GAAATGCTAAA	AGCAGATCGT	CCAGTGCCCT	TTTCAGTGCT	ACCGGCCTCT	CACCAAGCAG	2220
	TCCCTCCATGT	GAGCAACCCC	GAGACAAAAA	TGCTAAGTGG	GATCAAGAGA	GCAGCACTCG	2280
	GAGAGGGTGT	TTGCCAGTCT	GAGTGTCCCG	CGTGCCCGC	CAACCCGCTT	CCTGACTGAC	2340
	CTGAGCAAGG	TCTTACTAAG	CAGTCCCATC	TCTGTGGGAG	GCATGCAACG	CGTGCAGGGA	2400
40	GTTCAGGTGC	CGGTCCGGCT	AGCCAGGCCCT	GGAGGCCCCC	CAGGCAGGAG	GCCGCCCAAA	2460
	GGCGGGGCCG	CGCTCTCGCA	GACTAGGGGC	TGGGGGCGGC	CACAGACGGC	CTCGAAACCA	2520
	CAGCCCTTAC	CCCAATCCCA	CGAGCCCGCG	CAACGAACCA	CAGGTGCTGG	GCTTTAGAGA	2580
	ACATGGGAAG	GCGGCCCCAG	ACCTGGCCGG	AACGCCTTTC	CCTCAGAGCC	AGGCCCCGGC	2640
	CCCCTCTGGG	AAGCTCATCT	TGCGAAGCTG	AGGGAGCTCA	GGGCAAGGCC	CAGGCTAGCG	2700
45	CGGACCCGAA	GGGGCCGAGG	CTGCACGGGC	CTCTGCCAGA	ACGCTCAGGA	CATCCCGGCC	2760
	TGGGTTTACA	ACGCTGTTAG	GAAAATTAA	CAATGAATAA	AGCAACGTTT	AGTGCAGCA	

SEQ ID NO:267 PB9 Protein sequence:

Protein Accession #: NP_036561

MSGRVGDLSR RQKEALAKFR ENVQDVLPAI PNPDDYFLR WLRARSFDLQ KSEAMLRKHV 60
 EFRKQKIDIN IISWQPEVI QYLSGGMCG YDLGDCPVWY DIIGPLDAKG LLFSASKQDL 120
 LRTKMRCEL LLQECARHT KLGRKVEIT IYDCEGLGL KHLWKPAVEA YGEFLCMFEE 180
 NYPETLKRFL VVKAPKLPV AYNLIKPLS EDTRKKIMVL GANWKEVLLK HISPDQVPVE 240
 YGGTMDPDG NPKCKSKIN YGGDIPRYV RDQVKQYEH SVQISRSSH QVEYELFPG 300
 CVLRWQMSD GADVGFGLF KTKMGERQRA GEMTEVLNPN RYNHSLVPED GTLTCSDPGI 360
 YVLRFDNTYS FIHAKVNFT VEVLPLDKAS EEMKMKQLGAG TPK

SEQ ID NO:268 PBH8 DNA sequence

Nucleic Acid Accession#: XM_009756

Coding sequence: 301-1440 (underlined sequence corresponds to start and stop codon)

65	1	11	21	31	41	51	
	GTGGGGACAG	CCGAGCCGCG	CCGGGGCCCT	GGACGGCGTC	GCCAAGGAGC	TGGGATCGCA	60
	CTTGCTGCAG	ACTTTGGATG	GATTGTGTTT	TGTGGTAGCA	TCTGATGGCA	AAATCATGTA	120
	TATATCCGAG	ACCCTTCTCTG	TCCATTTAGG	CTTATCCCG	GTGGAGCTCA	CGGGCAACAG	180
	TATTTATGAA	TACATCCATC	CTTCTGACCA	CGATGAGATG	ACCGCTGTCC	TCACGGCCCA	240
70	CCAGCCGCTG	CACCACCACC	TGCTCCAAGG	TATGAGATAG	AGAGGTCTGT	CTTCTCTCGA	300
	ATGAAATGTG	TCTTGGCGAA	AAGGAACGCG	GGCCTGACCT	GCAGCGGATA	CAAGGTTCATC	360
	CACCTGAGTG	GCTACTTGAA	GATCAGGCAG	TATATGCTGG	ACATGTCCCT	GTACGACTCC	420
	TGCTACCAAG	TTGTGGGGCT	GGTGGCCGTG	GGCCAGTCGC	TGCCACCCAG	TGCCATCACC	480
	GAGATCAAGC	TGTACAGTAA	CATGTTTCATG	TTCAGGGCCA	GCCTTGACCT	GAAGCTGATA	540
75	TTCTTGGATT	CCAGGGTGAC	CGAGGTGAGC	GGGTACGAGC	CGCAGGACCT	GATCGAGAAG	600
	ACCTTATACC	ATCAGCTGCA	CGGCTGCGAC	GTGTTCACAC	TCCGCTACGC	ACACCACCTC	660
	CTGPTGGTGA	AGGGCCAGGT	CACCACCAAG	TACTACCGGC	TGCTGTCCAA	GCGGGGCGGC	720
	TGGTGTGGG	TGCAGAGCTA	CGCCACCGTG	GTGCACAACA	GCCGCTCGTC	CCGGCCCCAC	780
	TGCATCGTGA	GTGTCAATTA	TGTACTCACG	GAGATTGAAT	ACAAGGAAC	TCAGCTGTCC	840
80	CTGGAGCAGG	TGTCCACTGC	CAAGTCCCAG	GACTCTCGGA	GGACCCGCTT	GTCTACCTCA	900

CAAGAACTA GGAATAGT GAAACCCAAA AATACCAAGA TGAAGACAAA GCTGAGAACA 960
 AACCCTTACC CCCACAGCA ATACAGCTCG TTCCAAATGG ACRAACTGGA ATGCGGCCAG 1020
 CTCGGAACT GGAGAGCCAG TCCCCCTGCA AGCGCTGCTG CTCTCCAGA ACTGCAGCCC 1080
 CACTCAGAAA GCAGTGACCT TCTGTACACG CCATCCTACA GCCTGCCCTT CTCTACCAT 1140
 TACGGACACT TCCCTCTGGA CTCTCAGCTG TTCAGCAGCA AAAAGCCCAAT GTTGCCGGCC 1200
 AAGTTCGGGC AGCCCAAGG ATCCCTTGT GAGGTGGCAC GCTTTTCTT GAGCAACTG 1260
 CCAGCCAGCG GTGAATGCCA GTGGCATTAT GCCAACCCCT TAGTGCCTAG CAGCTCGTCT 1320
 CCAGCTAAAA ATCTCCAGA GCCACCGCG AACACTGCTA GGCACAGCCT GGTGCCAAGC 1380
 TACGAAGGCA AGCAGATGTC CTCTCGGAG ATACCGCCAG CTCCCCAGGA CGCAGACTGA 1440
 CTCCTGTTTG CTCGTGGAC CAAC

SEQ ID NO:269 PBH8 Protein sequence:

Protein Accession #: NP_005060

MKEKSKNAAK TRREKENGFE YELAKLLPLP SAITSQLDKA SIIRLTTSYL KMRAVFPEGL 60
 GDAWGPSPRA GPLDGVAKEL GSHLLQTLDG FVFVVASDVG IMYISETASV HLGLSQVELT 120
 GNSIYEIYHP SDHDEMTAVL TAHQPLHHHL LQYEIERSF FLRMKCVLAK RNAGLTCSGY 180
 KVIHCSGYLY IRQYMLDMSL YDSCYQIVGL VAVGQSLPPS AITEIKLYSN MFMFRASLDL 240
 KLIFLDSRVY EVTGEPQDL IEKTLYHHVH GCDVFHLRYA HHLLLVKGQV TKYYRLLSK 300
 RGGWVWVQSY ATVVHNSRSS RPHCIVSVNY VLTEIEYKEL QLSLEQVSTA KSQDSWRTAL 360
 STSQETRLKV KPKNTKMKTK LRTNPYPPOQ YSSFQMDKLE CGQLGNWRAS PPASAAAPPE 420
 LQPHSESSDL LYTPSYSLPF SYHYGHFPLD SHVFSSKKPM LPAKFGQPQG SPCEVARFFL 480
 STLPAEGECQ WHYANPLVPS SSSPAKNPPE PPANTARHSL VPSYEAPAAA VRRFGEDTAP 540
 PSFPCGHYR EEPALGPAKA ARQAARDGAR LALARAPEC CAPPTPEAPG APAQLPFVLL 600
 NYHRVLARRG PLGGAAPAAS GLACAPGPE AATGALRLRH PSPAATSPPG APLPHYLGA 660
 VIITNGR

SEQ ID NO:270 PBJ9 DNA sequence:

Nucleic Acid Accession#: AA760894

GGCACGAGGA GAAGATGTGG CTTGCTCATG CTTGACTTCT GCCATGGTTG TGAGGCCTCC 60
 CCAGCCATGT GGAAGCTGTT TCAGGTGCTG GTTCCATGGC TCTCTGAG CCGAAAATAA 120
 GGAAACTCCA TAGACCTTGT CCACTGGAAC TCGTTCCCAT CTACCCTCCA CTCTATCCAG 180
 GGTGATGGAT CTCTGCAGTA AGTGAAGAG TTTCTCATGG CCCCCAAGGT TATATCCATC 240
 TAGAACTTCA GCACGTAATT TCATCTGGAA ATAGTGCCTT TGTGGATATA AGTTAGGTAA 300
 AACTGAAGAT GAGATCATA TGGATTAGGA TGGGATCTAA ATCCAATGAA AATGTCTTCA 360
 TAAAAAACAG GAAAGAACCC ATAGAAACAC AAGGAAGAAG GTCATGTGAA GATGGAGGCA 420
 GAGATTGGAG GGATGCAGCC ACCGGCCCCAG GAATGCCAGC AGCCACCCAG AAGCTGGAAG 480
 GAAATGAGGG ATTTCTCTCT AGAACCTTTA GAGAGACAT GGTCTGTGTA ACAGCTTGAT 540
 TTTGGACTTG CCCATAGCTT GTATACTCTT ACTTTGGATA CAATTTTATC CAAACTTGGC 600
 TAAACAGTTT CTCAGCCTAT GGAAAATTTA AAATGGAGAA GATTCAACTC GATTTCTTACA 660
 GATTCAAAGC AAGAAAATGA TGGGAACATA GGAGGAGACC AAGAAAGCCT ATAAAAAGCA 720
 AAAATATGAA GTGAACATTG TGGTAGCTTT AAGATGTGTA GTGTAGCTGC AGGCACCCCTA 780
 TACACATGAA AACCCCAAG GGAATCCCC ATATCACAGT GTAGTGTGAT ATTGACATT 840
 YGTGATCATY TAGAGATGTA CAGAAAAGGT GAATCTGTGT TCTGTATATT CTGCCTAAGG 900
 CAAAGAAATG TTTAGCTTTC TTTAAATAG TTCCATAATT TTTTYAAAA AGCTTTGCTT 960
 GAAAACTGTA AGCTTCCCAT ATCTGGAGCA TTTCACTTTA AATATTGGA TAAATATGTT 1020
 ATCTTCTTAC TTGGACATTT CATGTGTTA GGGATTGTYT TYTAAATCTT TCCTAATTCA 1080
 TATAGCTGCT AACACTTCCC GCAGAGCTAA ACCATTACAG ANTATGAAAT AAAGACCCTA 1140
 TTGATTGAA CTAAAAAAA AAAMAMAAAA AAAAAAAAAA AAAAAAAAT GA

SEQ ID NO:271 PBQ4 DNA sequence

Nucleic Acid Accession#: AA149579

Coding sequence: 1-1363 (underlined sequence corresponds to start and stop codon)

1 11 21 31 41 51
 | | | | |
 60 ATGGAATCAA TCTCTATGAT GGAAGCCCT AAGAGCCTTA GTGAAACTTG TTTACCTAAT 60
 GGCATAAATG GTATCAAAGA TGCAAGGAAG GTCACGTGAG GTGTGATTGG AAGTGGAGAT 120
 TTTGCCAAAT CCTTGACCAT TCGACTTATT AGATGCGGCT ATCATGTGGT CATAGGAAGT 180
 AGAAATCCTA AGTTTGCTTC TGAATTTTCT CCTCATGTGG TAGATGTCAC TCATCATGAA 240
 65 GATGCTCTCA CAAAAACAAA TATAATATTT GTTGCTATAC ACAGAGAACA TTATACCTCC 300
 CTGTGGGACC TGAGACATCT GCTTGTGGGT AAAATCCTGA TTGATGTGAG CAATAACATG 360
 AGGATAAACCA AGTACCAGCA ATCCAATGCT GAATATTGCG CTTCATTATT CCCAGATTCT 420
 TTGATTGTCA AAGGATTATA TGTGTCTCA GCTTGGGCAC TTCAGTTAGG ACCTAAGGAT 480
 GCCAGCCGCG AGGTTTATAT ATGCAGCAAC AATATTCAAG CGCACAACA GGTATTGAA 540
 70 CTTGCCCGCC AGTTGAATTT CATTCCCATT GACTTGGGAT CCTTATCATC AGCCAGAGAG 600
 ATTGAAATTT TACCCCTACG ACTCTTACT CTCTGGAGAG GGCCAGTGGT GGTAGCTATA 660
 AGCTTGGCCA CATTTTCTTT CTTTATTCCT TTTGTACAG ATGTGATTCA TCCATATGCT 720
 AGAAACCAAC AGAGTACTT TTACAAATTT CCTATAGAGA TGTGAATAA AACCTTACCT 780
 ATAGTTGCCA TTACTTTGCT CTCCCTAGTA TACCTCGCAG GTCTTCTGGC AGCTGCTTAT 840
 75 CAACCTTATT ACAGCAACCA GTATAGGAGA TTTCACCTT GGTGGAAAC CTGGTTACAG 900
 TGTAAGAAAC AGCTTGGATT ACTAAGTTT TTCTTCGCTA TGGTCCATGT TGCCCTACAGC 960
 CTCTGCTTAC CGATGAGAAG GTCAGAGAGA TATTGTTTC TCAACATGGC TTATCAGCAG 1020
 GTTCATGCAA ATATTGAAAA CTCTTGGAAAT GAGGAAGAAG TTTGGAGAAT TGAATGTAT 1080
 ATCTCCTTTG GCATATAGAG CTTTGGCTTA CTTCCTCTCC TGGCAGTCAC TTCTATCCCT 1140
 TCAGTGAGCA ATGCTTTAAA CTGGAGAGAA TTCAGTTTTC TTCAGTCTAC ACTTGATAT 1200

GTCGCTCTGC TCATAAGTAC TTTCCATGTT TTAATTTATG GATGGAAACG AGCTTTTGAG 1260
 GAAGAGTACT ACAGATTTTA TACACCACCA AACTTTGTTC TTGCTCTTGT TTTGCCCTCA 1320
 ATGTGAATTC TGGATCTTTT GCAGCTTTGC AGATACCCAG ACTGA

5 SEQ ID NO:272 PBQ4 Protein sequence:

Protein Accession #: none

10 1 11 21 31 41 51
 MESISMMGSP KSLSETCLPN GINGIKDARK VTVGVIGSGD FAKSLTIRLI RCGYHVIVGS 60
 RNPKFASEFF PHVVDVTHHE DALTKTNIIF VAIHREHYTS LWDLRHLLVG KILIDVSNM 120
 RINQYPESNA EYLASLFPDS LIVKGFNVVS AWALQLGPKD ASRQVYICSN NIQARQQVIE 180
 LRQLNFIPI DLGSLSSARE IENLPLRLFT LWRGPFVVVAI SLATFFFLYS FVRDVIHPYA 240
 15 RNQSDIFYKI PIEIVNKTLP IVAITLLSLV YLAGLLAAAY QLYYGTKYRR FPPWLETWLQ 300
 CRKQLGLLSF FFMAMHVAYS LCLPMRRSER YLFLNMAYQQ VHANIENSWN EEEVWRIEMY 360
 ISFGIMSLGL LSLAVTSIP SVSNALNWRE FSFIQSTLGY VALLISTPHV LIYGWKRAFE 420
 EEEYRYFYTPP NFVLALVLPS IVILDLLQLC RYPD

20 SEQ ID NO:273 PBQ5 DNA SEQUENCE

Nucleic Acid Accession#: NM_001973

Coding sequence: 150-1445 (underlined sequence corresponds to start and stop codon)

25 1 11 21 31 41 51
 CGCGCCGCTT CTACTCCGCC GCGGGGGTCG CAGCGGCTGC CGCGCGCTCC TCGAGTTTCC 60
 AGCGTGAGGA GGAGGCTGAG GCGGAGAGG CGCATCGTGT TCGAGGCGGA GACCGAGGGG 120
 GAGCCCCCGC CGCGGCGTCG CTCATTGCTA TGGACAGTGC TATCACCCCTG TGGCAGTTCC 180
 30 TTCTTCAGCT CCTGCAGAAG CCTCAGAACA AGCACATGAT CTGTTGGACC TCTAATGATG 240
 GGCGAGTTAA GCTTTTGCAG GCAGAAGAGG TGGCTCGTCT CTGGGGGATT CGCAAGAACA 300
 AGCCTAACAT GAATTTATGAC AACTCAGCC GAGCCCTCAG ATACTATTAT GTAAAGAATA 360
 TCATCAAAAA AGTGAATGGT CAGAAGTTTG TGTACAAGTT TGTCTCTTAT CCAGAGATT 420
 TGAACATGGA TCCAATGACA GTGGGCGAGGA TTGAGGGTGA CTGTGAAAGT TTAAACTTCA 480
 GTGAAGTCAG CAGCAGTTCC AAAGATGTGG AGAATGGAGG GAAAGATAAA CCACCTCAGC 540
 35 CTGGTGCCAA GAGCTCTAGC CGCAATGACT ACATACACTC TGGCTTATAT TCTTCATTTA 600
 CTCTCAACTC TTTGAACTCC TCCAATGTAA AGCTTTTCAA ATTGATAAAG ACTGAGAATC 660
 CAGCGGAGAA ACTGGCAGAG AAAAAATCTC CTCAGGAGCC CACACCATCT GTCATCAAA 720
 TTGTACAGAC ACCTTCCAAA AAGCCACCCG TTGAACCTGT TGCTGCCACC ATTTCAATTG 780
 40 GCCCAAGTAT TTTCTCATCT TCAGAAGAAA CTATCCAAGC TTTGGAGACA TTGGTTTCCC 840
 CAAAAC TGCC TTTCCCTGGAA GCCCAACCT CTGCCTCTAA CGTAATGACT GCTTTTGCCA 900
 CCACACCACC CATTTCTGTC ATACCCCTTT TGCAGGAACC TCCCAGAACA CCTTCACCAC 960
 CACTGAGTTC TCACCCAGAC ATCGACACAG ACATGTATTC AGTGGCTTCT CAGCCCAATGG 1020
 AACTTCCAGA GAATTTGTCT CTGGAGCCTA AAGACCAGGA TTCAGTCTTG CTAGAAAAGG 1080
 45 ACAGAAGTAA TAATTCATCA AGATCCAAGA ACCCAAAGG GTTAGGACTG GCACCCACCC 1140
 TTGTGATCAC GAGCAGTGAT CCAAGCCAC TGGGAATACT GAGCCCATCT CTCCCTACAG 1200
 CTCTCTTAC ACCAGCATTT TTTTCACAGA CACCCATCAT ACTGACTCCA AGCCCTTGC 1260
 TCTCCAGTAT CCACTTCTGG AGTACTCTCA GTCTGTGTC TCCCCTAAGT CCAGCCAGAC 1320
 TGCAAGGTGC TAACACACTT TTCCAGTTTC CTCTGTACT GAACAGTCAT GGGCCATTCA 1380
 50 CTCTGTCTGG GCTGGATGGA CCTTCCACCC CTGGCCCAT TTTCCAGAC CTACAGAAGA 1440
 CATTAACCTAT GCACCTGTGG AATGAGAGAA CCGAGGAACG AAGAAACAGA CATTCACAT 1500
 GATTGCAATTT GAAGTGAACA ATGTAGATT CTACAATGCT GATAATAGAC TATTGTGATT 1560
 TTTGCCATT CCACTGAAA ACATCTTTT AGGATTCTCT TTGAATAGGA CTCAGTTGG 1620
 ACTATATGTA TAAAAATGCC TTAATTGGAG TCTAACTCC ACCTCCCTCT GTCCTTTTCT 1680
 55 TTTCTTTTTC TTTCTTCTCT TCTTTTCTT TTTCTCTTCA AAAATATTTT GAGCTTTGTG 1740
 CTGAAGAAGT TTTTGTGGG CTTTGTGAC TGTGCTTTC AAAAGCAATT AAGAACAAAG 1800
 TTAATCTCTC TGCTATGAG GACCTTTTG CCAGGAAAAA TTATGCTTAG AATCTATTAT 1860
 TTAAGAAGT ATTTGTGAAA TGAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA 1920
 AAAAAAAA AAA

60 SEQ ID NO:274 PBQ5 Protein sequence:

Protein Accession #: NP_001964

65 MDSAITLWQF LLQLLQKPQN KHMICWTSND GQFKLLQAE VARLWGIRKN KPNMNYDKLS 60
 RALRYYYVKN IIKVNGQKF VYKFVSYPEI LNMDPMTVGR IEGDCESLNF SEVSSSSKDV 120
 ENGGKDKPPQ PGAKTSSRND YHSGLYSSF TLNSLNSSNV KLFKLIKTN PAEKLAEKKS 180
 PQEPTPSVIK FVTTPSKKPP VEPVAATISI GPSISPSSEE TIQALETLSV PKLPSLEAPT 240
 SASNVMTAFA TTPPISSIPP LQEPRTSPS PLSSHPDIDT DIDSVASQPM ELPENLSLEP 300
 70 KDQDSVLEK DKNVNSSRSK KPKGLGLAPT LVITSSDSP LGILSFLPT ASLTPAFFSQ 360
 TPILTPSPL LSSIHFWSTL SPVAPLSPAR LQGANLTFQF PSVLNSHGPF TLGLDGPST 420
 PGFFSFDLQK T

75 SEQ ID NO:275 PBQ3 DNA SEQUENCE

Nucleic Acid Accession#: AB040921

Coding sequence: 131-2560 (underlined sequence corresponds to start and stop codon)

1 11 21 31 41 51

5
10
15
20
25
30
35
40
45
50

AATCAGGAAC AGATCATATA TTGACCGAGA TTCTGAGTAT CTCTTGCAAG AAAATGAACC 60
 AGATGGAAC TTAGACCAAA AATTATTGGA AGATTTACAA AAGAAAAAA ATGACCTTCG 120
 GTATATTGAA ATGCAGCATT TCAGAGAAAA GCTGCCTTCG TATGGAATGC AAAAGGAATT 180
 GGTAATTTA ATTGATAACC ATCAGGTAAC AGTAATAAGT GGTGAACTG GTTGTGGCAA 240
 AACCACCTCA GTTACTCAGT TCATTTTGGG TAACATACAT GAAAGAGGAA AAGGATCTGC 300
 TTGCAGAAAT GTTTGTACTC AGCCAAGAAG AATTAGTGCC ATTTTCAGTTG CGGAAAGAGT 360
 AGCTGCAGAA AGGGCAGAAT CTTGTGGCAG TGGTAATAGT ACTGGATATC AAATTCGTCT 420
 CCAGAGTCGG TTGCCAAGGA AACAGGGTTC TATCTTATAC TGTACACAG GAATCATCCT 480
 TCAGTGGCTC CAGTCAGACC CGTATTGTG CAGTGTAGT CATATCGTAC TTGATGAAAT 540
 CCATGAAAGA AATCTGCAGT CAGATGTTT AATGACTGTT GTTAAAGACC TTCTCAATTT 600
 TCGATCTGAC TTGAAAGTAA TATTGATGAG TGCAACATG AATGCAGAAA AGTTTTCAGA 660
 ATATTTTGGT AACTGTCCAA TGATACATAT ACCTGGTTTT ACCTTTCCGG TTGTGGAATA 720
 TCTTTTGGAA GATGTAATTG AAAAAATAAG GTATGTTCCA GAACAAAAAG AACACAGATC 780
 CCAGTTTAAG AGGGGTTTCA TGCAAGGGCA TGTAATAGA CAAGAAAAAG AAGAAAAAGA 840
 AGCAATATAT AAAGAACGTT GGCCAGATTA TGTAAGGGAA CTGCGAAGAA GGTATTCTGC 900
 AAGTACTGTA GATGTTATAG AAATGATGGA GGATGATAAA GTTGATCTGA ATTTGATTTG 960
 TGCCCTCATC CGATACATTG TTTTGGAAAG AGAGGATGGT GCGATACCTG TCTTCTGCC 1020
 AGGCTGGGAC AATATCAGCA CTTTACATGA TCTCTTGATG TCACAAGTAA TGTTTAAATC 1080
 AGATAAATTT TTAATTTATC CTTTACATTC ACTGATGCCT ACAGTTAACC AGACACAGGT 1140
 GTTTAAAAAG ACCCTCTCCT GTGTTCCGAA AATAGTAATT GCTACCAACA TTGCGGAGAC 1200
 TAGCATTACC ATAGATGATG TCGTTTATGT GATAGATGGA GGAAAAATAA AAGAGACGCA 1260
 TTTTGATACT CAGAACAATA TCAGTACAAT GTCCGCTGAG TGGGTTAGTA AAGCTAATGC 1320
 CAAACAGAGA AAAGGTCGAG CTGGAAGAGT TCAACCTGGT CATTCGTATC ATCTGTATAA 1380
 TGGTCTTAGA GCAAGTCTTC TAGATGACTA TCAACTGCCA GAAATTTTGA GAACTCCTTT 1440
 GGAAGAACTT GTTTTACAAA TAAAGATTTT AAGGCTAGGT GGAATTTGCTT ATTTTCTGAG 1500
 TAGATTAATG GACCCACCAT CAAATGAGGC AGTGTACTC TCCATAAGAC ACCTGATGGA 1560
 GCTGAACGCT TTGGATAAAC AAGAAGAATT GACACCTCTT GGAGTCCACT TGGCACGATT 1620
 ACCCGTTGAG CCACATATG GAAAAATGAT TCTTTTGGG GCACTGTCTT GCTGCTTAGA 1680
 CCCAGTACTC ACTATTTGCTG CTAGTCTCAG TTTCAAAGAT CCATTTGTC TCCACTGGG 1740
 AAAAGAAAAA ATTGAGATG CAAGAAGAAA GGAATTGGCA AAGGATATAA GAAGTATGTA 1800
 CTTAACAGTT GTGAATGCGT TTGAGGGCTG GGAAGAGGCT AGGCGACGTG GTTTCAGATA 1860
 CGAAAAAGAC TATTGCTGGG AATATTTTCT GTCTTCAAAC AACTGTCAGA TGCTGCATAA 1920
 CATGAAAGGA CAGTTTCTG AGCATCTTCT TGGAGCTGGA TTTGTAAGCA GTAGAAATCC 1980
 TAAAGATCCA GAATCTAATA TAAATTCAGA TAATGAGAAG ATAATTAAAG CTGTCATCTG 2040
 TGCTGTTTAT TATCCCAAAG TTGCTAAAT TCGACTAAT TTGGGTAAAA AAAGAAAAAT 2100
 GGTAAGAGTT TACACAAAA CCGATGGCCT GGTGTCTGTT CATCTAAAT CTGTTAATGT 2160
 GGAGCAAAAC GACTTTTCT ACAAATGGCT TATCTATCAC CTAAAGATGA GAACAAGCAG 2220
 TATATACTTG TATGACTGCA CAGAGGTTTC CCCATACTGT CTCTTGTTTT TTGGAGGTGA 2280
 CATTTCCATC CAGAAGGATA ACGATCAGGA AACTATGCT GTAGATGAGT GGATTTGATT 2340
 TCAGTCTCCA GCAAGAATTG CCCATCTTGT TAAGGAATTA AGAAAGGAAC TAGATATTCT 2400
 TCTGCAAGAG AAGATTGAAA GTCTCTATCC TGTAGACTGG AATGACACTA AATCCAGAGA 2460
 CTGTGCAGTA CTGTACAGTA TTATAGACTT GATCAAAACA CAGGAAAGG CAACCTCCAG 2520
 GAACTTTCCG CCACAGATCC AGGATGGATA TTACAGCTGA CAGCTTTTCA GGGGTGGTCT 2580
 GAAAAGCCAG TTTGACAGCC ATTCTTCATC ATTTGTTTAAA TTTTGGCTGG ATGCCAAACC 2640
 CTGGGACATG AACAATTTTC ATGTGTAAGG TAGAAGCCTT CAGTAGGTAG TAAAGACTTA 2700
 ATGTGCATGA CTTGATGTTA TATGTAGAGA TATATATATA TATATATATA CCATAAAAGC 2760
 AATATGTTCT CTGATCATAT ACTCTGCTGT GGTATGCCCC ACTCTTTGGG AGTATATATC 2820
 CTTTATATAT ATTGATATT GTACCACTTG AGAAATTCCT TTGTCTGTTT ATACAAAAT 2880
 AATCTTTCTG CTCATAATGA TTGATGATAC CACCAGTAAA AATAGGATGT TTACCCCAA 2940
 ACAAGTGTA ATTAAGAAAT TGAACACAAC CACATTTTAT AAAATGAAC TTCTATCGGA 3000
 AGTAATTA TTTGTGTGTA TAAAGTCCAG TATTTAATAA AATGTACAAT GTTAAATCTC

SEQ ID NO:276 PB3 Protein sequence:

Protein Accession #: BAA96012

55
60
65
70

IRNRSYIDRD SEYLLQENEP DGTLDQKLE DLQKKKNDLR YIEMQHFREK LPSYGMQKEL 60
 VNLDNHQVT VISGETGCGK TTQVTQFILD NYIERGKGS CRIVCTQPRR ISASVAERV 120
 AAERAESCGS GNSTGYQRL QSRLPRKQGS ILYCTTGIL QWLQSDPYLS SVSHVLDL 180
 HERNLQSDVL MTVVKDLNLF RSDLKVLMS ATLNAEFSE YFGNCPMIHI PGFTFPVVEY 240
 LLEDVIEKIR YVPEQKEHRS QFGRGFMQGH VNRQEKKEKE AIYKERWPDY VRELRRRYSA 300
 STDVVIEMME DDKVDNLNLI ALIRYIVLEE EDGAILVFLP GWDNISTLHD LLMSQVMFMS 360
 DKFLIPLHS LMPVTNQTV FKRTPPGVRK IVIATNIAET SITIDDVVYV IDGGKIKETH 420
 FDTQNNISM SAEWVSKANA QQRKGRAGRV QPGHCYHLYN GLRASLLDDY QLPEILRTP 480
 EELCLQIKIL RLGGIAYFLS RLMDPPSNEA VLLSIRHLE LNALDKQEEL TPLGVHLARL 540
 PVEPHIGKMI LFGALFCCLD PVLTAASLS FKDPFVIPLG KEKIDARRK ELAKDTRSDH 600
 LTVVNAFEGW BEARRRGFRY EKDYCWEYFL SSNTLQMLHN MKGQFAEHL GAGFVSSRN 660
 KDPESNINS NEKIKAVIC AGLYPKVAKI RLNLGKKRKM VKVYTKIDGL VAVHPKSVNV 720
 EQTDFHYNWL IYHLKMTSS IYLYDCTEVS PYCLLFFGGD ISIQKNDQE TIAVDEWIVF 780
 QSPARIAHLV KELRKELDIL LQEKIESPHV VDWNDRKSRD CAVLSAIDL IKTQEKATPR 840
 NFPPRFQDGY YS

SEQ ID NO:277 PB6 DNA SEQUENCE

Nucleic Acid Accession#: AA464018

Coding sequence: 64-1669(underlined sequence corresponds to start and stop codon)

GATTTTATCC TGGAAACATTA CAGTGAAGAT GGCTATTTAT ATGAAGATGA AATTGCAGAT 60
 CTTATGATC TGAGACAAGC TTGTCGGACG CCTAGCCGGG ATGAGGCCGG GGTGGAAC 120

CTGATGACAT ACTTCATCCA GCTGGGCTTT GTCGAGAGTC GATTCTTCCC GCCACACGG 180
 CAGATGGGAC TCCTGTTCAC CTGGTATGAC TCTCTACCG GGGTTCGGT CAGCCAGCAG 240
 AACCTGCTGC TGGAGAAGGC CAGTGTCTG TTCAACACTG GGGCCCTCTA CACCCAGATT 300
 5 GGGACCCGGT GTGATCGGCA GACGCAGGCT GGGCTGGAGA GTGCCATAGA TGCCTTTCAG 360
 AGAGCCCGAG GGGTTTTAAA TTACCTGAAA GACACATTTA CCCATACTCC AAGTTACGAC 420
 ATGAGCCCTG CCATGCTCAG CGTGCTCGTC AAAATGATGC TTGCACAAGC CCAAGAAAGC 480
 GTGTTTGAGA AAATCAGCCT TCCTGGGATC CGGAATGAAT TCITCATGCT GGTGAAGGTG 540
 10 GCTCAGGAGG CTGCTAAGGT GGGAGAGGTC TACCAACAGC TACACGCAGC CATGAGCCAG 600
 GAGCCCGGTGA CACCTTGCG CACACTGAAG AATGATCAGC AGCGCCGACA GCTGGGGAAG 660
 CACTACGGGG CCCTGGCCCA CTACTTCACT GCCATCTCC TCATCGACCA CCAGGTGAAG 720
 CCAGGCACGG ATCTGGACCA CCAGGAGAAG TGCCTGTCCC AGCTCTACGA CCATATGCCA 780
 15 GAGGGGCTGA CACCTTGCG CACACTGAAG AATGATCAGC AGCGCCGACA GCTGGGGAAG 840
 TCCCACTTGC GCAGAGCCAT GGCTCATCAC GAGGAGTCGG TCGGGGAGGC CAGCCTCTGC 900
 AAGAAGCTGC GGAGCATTGA GGTGCTACAG AAGGTGCTGT GTGCCGCACA GGAACGCTCC 960
 20 GAGCTCAGCT ACGCCGACA CCAGGAGGAG GATGACCTGC TGAACCTGAT CGACGCCCCC 1020
 AGTGTGTGTG CTAACACTGA GCAAGAGGTT GACATTATAT TGCCCCAGIT CTCCAAGCTG 1080
 ACAGTCACGG ACTTCTTCCA GAAGCTGGGC CCCTTATCTG TGTTCGCG TAACAAGCGG 1140
 TGGACGGCTC CTCGAAGCAT CCGCTTCACT GCAGAAGAAG GGGACTTGGG GTTCACCTTG 1200
 AGAGGGAACG CCCCCGTTA GGTTCACCTC CTGGATCCTT ACTGCTCTGC CTCGGTGGCA 1260
 25 PGTDLDHQEK CLSGLYDHMP EGLTPLATLK NDQRRQLGK SHLRRAMAHH EESVREASLC 1320
 CTGAGTGAGG TTATGAAGCT GCTGAAGAGC TTTGGCGAGG ACGAGATCGA GATGAAAGTC 1380
 GTGAGCCTCC TGGACTCCAC ATCATCCATG CATAATAAGA GTGCCACATA CTCCGTGGGA 1440
 ATGCAGAAAA CGTACTCCAT GATCTGCTTA GCCATTGATG ATGACGACAA AACTGATAAA 1500
 30 AACAGAAAA TCCTCAAGAA GCTTTCCTTC CTGAGTTGGG GCACCAACAA GAACAGACAG 1560
 AAGTCAGCCA GCACTTGTG CCTCCATCG CTCGGGGCTG CACGGCTCA GGTCAAGAAG 1620
 AAGCTGCCCT CCCCTTCAG CCTTCTCAAC TCAGACAGTT CTGAGTACTA A

SEQ ID NO:278 PBYS Protein sequence:
 Protein Accession #: NP_149094

DFILEHSEYD GYLYEDEIAD LMDLRQACRT PSRDEAGVEL LMTYFIQLGF VESRFFPPTR 60
 QMGLLFTWYD SLTGVPVSQQ NLLLEKASVL FNTGALYTOI GTRCDROTQA GLESAIDAFQ 120
 35 RAAGVLNLYLK DITHTPSYD MSPAMLSVLV KMMLAQAES VFEKISLPGI RNEFFMLVKV 180
 AQEAARKVEV YQQLHAAMSQ APVKENIPYS WASLACVKAH HYAALAHYFT AILLIDHQVK 240
 PGTDLDHQEK CLSGLYDHMP EGLTPLATLK NDQRRQLGK SHLRRAMAHH EESVREASLC 300
 KKLRSIEVLQ KVLCAAQERS RLTYAQHQEE DDLNLDIAP SVVAKTEQEV DILPQFSKL 360
 TVTDFQKLG PLSVFSANKR WTPPRSIRFT AEEGDLGFTL RGNAPVQVHF LDPYCSASVA 420
 40 GAREGDYIVS IQLVDCKWLT LSEVMKLLKS FGEDEIEMKV VSLDSTSSM HNKSATYSVG 480
 MQKTYSMICL AIDDDDKTDK TKKISKLSF LSWGTNKNRQ KSASTLCLPS VGAARPQVKK 540
 KLSPFSLN SDSSWY

SEQ ID NO:279 PBYS DNA SEQUENCE

Nucleic Acid Accession#: AF107493
 Coding sequence: 125-556 (underlined sequence corresponds to start and stop codon)

	1	11	21	31	41	51	
50	GAATTCGGCA	CGAGCCTTGT	TGGAGGTTCT	GGGGCGCAGA	ACCGCTACTG	CTGCTTCGGT	60
	CTCTCCTTGG	GAATAAATAA	AAATTGAACC	TTTGTGGAGCT	GTGTGCTAAA	TCTTCAGTGG	120
	GACAATGGGT	TCAGACAAAA	GAGTGAAGTAG	AACAGAGCGT	AGTGGAAAGAT	ACGGTTCCAT	180
	CATAGACAGG	GATGACCCGT	ATGAGCGTGA	ATCCCGAAGC	AGCGCGAGGG	ACTCAGATTA	240
	CAAAAGATCT	AGTGATGATC	GGAGGGGTGA	TAGATATGAT	GACTACCGAG	ACTATGACAG	300
55	TCCAGAGAGA	GAGCGTGAAA	GAAGGAACAG	TGACCGATCC	GAAGATGGCT	ACCATTCAGA	360
	TGGTGACTAT	GGTGAGCAGC	ACTATAGGCA	TGACATCAGT	GACGAGAGGG	AGAGCAAGAC	420
	CATCATGCTG	CGCGGCCCTC	CCATCACCAT	CACAGAGAGC	GATATTCGAG	AAATGATGGA	480
	GTCTTTCGAA	GGCCCTCAGC	CTGCGGATGT	GAGGCTGATG	AAGAGGAAAA	CAGGTGAGAG	540
	CTTGCTTAGT	TCCTGATATT	ATTGTTCTCT	TCCCCATTCC	CACCTCAGTC	CCTAAAGAAC	600
60	ATCCTGATTC	CCCCAGTCTT	CAAGCACATG	AATTCAGAAT	GAAAGGTTTG	CCATGGCTAA	660
	GGAAATGTAC	TCCTTGAATA	CCATGTTAGC	ATCTGAGGAA	CTTTTATAAA	CTTTGTTTAA	720
	GGGACTTTT	TTTCTTAGG	TAAGTAATGA	TTTATAAACT	CCTTTTATTT	TTTGACTATA	780
	GTCCGTTTGA	TGGTACTTT	AAGCGTGGAA	TCAAATGGAG	TGGCATTAG	TTCAGGCGGC	840
	TTGTTCCTTG	CCATGGCAAA	GTATCAAGAA	GATCCCCAAG	TCAAGTCACA	TTGTGTAAGC	900
65	TGCTTCCCAA	TTGGCTTTGT	CACGCAGTGT	TGAAGCAGTG	GGAGAGAGAT	TCACCTGTTA	960
	TAAAGGAAC	GACTAACACA	AGTATCCCGT	CTATATCTGA	ATGCTGTCTC	TAGGTGTAAG	1020
	CCGTGGTTTC	GCTTTCGTGG	AGTTTATCA	CTTGCAAGAT	GCTACCGACT	GGATGGAAGC	1080
	CAATCAGGTT	GCTTCACTCA	CCAAGTCTAG	ATATTCATGA	AAATGGAACA	AGTCTGTACA	1140
	ATTTTAAAAA	AAGGTTGAAG	GAGTGGTTTG	TTCCAAAGGA	GTGACTTTT	TTTAAAAAAA	1200
70	AAGCTTTGTA	TATATTAAAA	TTGATGTTAC	TAGAATAAGT	ACAGTACCAA	GGACTTCATT	1260
	ATAGAATTG	TTCTGCCTTT	AAACATGGCT	ACCTACCTGG	CAGGGCTTTG	TTAACTACTG	1320
	AATACCTGTC	TGGTAATCAC	TAAACATCT	TTATGTTTCC	CTTTTCTCTA	GTTTGTATTA	1380
	TTCTTATTAT	GTCTATTGAG	AGTAAGCTTA	GTATATCAAA	CTCTCCATTT	GACAGTGAAG	1440
	AGAACATAGT	GAAAGTCTGT	GGCGGCATTT	TTATAAGTAA	TTCTTTATTT	CTGCCTGAAG	1500
75	ACCACAAAGC	CTCCTGGAGG	CGTAACCTGT	CAGACCGGTC	TTCAGGGAAT	ATTTAAGGAC	1560
	TTAGTGGAA	TTATGAACAA	TAGTCTGAT	GAGATTAGCC	TGGGAGTGGT	GTCTGCAGC	1620
	TGCTTAATCT	AGAGTGGCAT	TAACATTCTA	ATCTCCTTGA	GAATGCCTTT	TATAGTCTGT	1680
	TCAAAGCAAG	TCATTGATGG	TTCTTCGAGG	TAGTGTAAAC	TGAAGTGTTC	TTCAAGTTGT	1740
	CAAGATAATG	TTCAAGTCTT	GGCACTTAAA	TAACATTTTT	TGCAAGAACT	CCAAGGCACA	1800

5 TTATTGAATG CCTTTAACCA AGTGCATCTT GGGAAAGTTG CTTGACTCAT TATCTTGCTT 1860
 TTCTGCAGCA TTCTGTGATT TGAGTCATCC ATGAATCCAT GAATAAAAGT TACATTCTTT 1920
 GATTGGTAAT ATTGCCATTT ATAACAAGAC TCACTAATGA GGGTATCACT TTGACTGACT 1980
 GATTGTGTTAA AGTTTTTAAG CCTCTCATTT TCCTAACCCA GAAATCACAG CCTGATTTTA 2040
 10 TTAAGTAG AGCTTCATTC ATTTCAATACC ATAGATACCA TCCTAGTAAA TCCAGAACAT 2100
 ATACAAGGTT CATGTGAGTC TGCTTTCTTG ACATGATAGC ATTTGTTGAT GCAGTGGATA 2160
 TGTCAGAAATG ACTAACCTAG GAGTTTGAAA CTCCTAAGAA ACTAAAACCT GTAAAGACATT 2220
 TAAAGTCTC CACAATTTTA ATGTATACAA AGCTATGTTA CTGTGTAACA CATTACAGTT 2280
 CAAATTCAC CCAGAAATAA AAGGCCAGTA GGATTAGGGA CTCACTGGTA GTTTGGAGTC 2340
 TCCCAGCACA CATCCCTCCT AGTGGGATGA TCTATTACAA TATCTCCCAG CTTTTTTTAT 2400
 TTTGCTTCTG TATATCACAG TGAGTGGATG GCCCTTCAGC TTTTCTCTC CTGGCCAGAC 2460
 ATGCAGTCTT GCCTTTAGAT ATCGCAGAGA CAAAATTCAC AGCATGTCTT AAATCTTCCA 2520
 GGATTTGCAA GAACCAAAAT GCTCAACAGT ATGTATGTTT AGAGGGGTTA GACTCCTTTT 2580
 15 TAAATCTGG ATATCTAACC ACCTACTTAA ATCTGTTTGA TAGTGTCAAA CCACCCCCAC 2640
 CCTGTATCCT CCCACCCCA AAAAAAAAAA AAAA

SEQ ID NO:280 PBV8 Protein sequence:

Protein Accession #: XP_003261

20 MGSDKRVSR ERSGRYGSII DRDDRDERES RSRRRSDSYK RSSDDRRGDR YDDYRDYDSP 60
 ERERERRNSD RSEDGYHSDG DYGEHDYRHD ISDERESKTI MLRGLPITIT ESDIREMMES 120
 FEGQPADVR LMKRKTGESL LSS

SEQ ID NO:281 PC12 DNA SEQUENCE

Nucleic Acid Accession#: AF208291

Coding sequence: 109-3705 (underlined sequence corresponds to start and stop codon)

30 1 11 21 31 41 51
 | | | | |
 CGGCCGCTTT TTTCTCAAGA TGGCAGATTC CCACTGAGGC TGAGGGGGCC GAGCTCGCGC 60
 GCCGCGTTCC CTCTCCGTTT GCCATGAACC GCGGACACCC CGGCCCCGAT GGCCCCCGTG 120
 35 TACGAAGGTA TGGCCTCACA TGTGCAAGTT TTCTCCCCCT ACACCCCTCA ATCAAGTGCC 180
 TTCTGTAGTG TGAAGAAACT AAAAGTAGAG CCAAGTTCCA ACTGGGACAT GACTGGGTAC 240
 GGTCTCCACA GCAAAGGTA CAGCCAGAGC AAGAACATAC CACCTTCTCA GCCAGCCTCC 300
 ACACCCGTCA GCACCTCCTT GCCGGTCCCA AACCCAGCC TACCTTACGA GCAGACCATC 360
 GTCTTCCAG GAAGCACCGG GCACATCGTG GTCACTCAG CAAGCAGCAC TTCTGTACCC 420
 40 GGGCAAGTCC TCGGCGGACC ACACAACCTA ATGCGTGGAA GCACTGTGAG CCTCCTTGAT 480
 ACCTACCAAA AATGTGGACT CAAGCGTAAG AGCGAGGAGA TCGAGAACAC AAGCAGCGTG 540
 CAGATCATCG AGGAGCATCC ACCCATGATT CAGAAATAAT CAAGCGGGGC CACTGTCCGC 600
 ACTGCCACCA CGTCTACTGC CACCTCCAAA AACAGCGGCT CCAACAGCGA GGGCGACTAT 660
 CAGCTGGTGC AGCATGAGGT GCTGTGCTCC ATGACCAACA CCTACGAGGT CTTAGAGTTC 720
 45 TTGGGCGGAG GACGTTTGGT ACAAGTGGTC AAGTGTGGA AACGGGGCAC CAATGAGATC 780
 GTAGCCATCA AGATCTTGAA GAACCGCCCA TCCTATGCCC GACAAGGTCA GATTGAAGTG 840
 AGCATCTGG CCGGCTTGAG CACGGAGAGT GCCGATGACT ATAACCTTCG TCGGGCTTAC 900
 GAATGCTTCC AGCACAAGAA CCACACGTGC TTGGTCTTCG AGATGTTGGA GCAGAACCCT 960
 TATGACTTTC TGAAGCAAAA CAAGTTTAGC CCCTTGCCCC TCAAAATACAT TCGCCAGTT 1020
 50 CTCCAGCAGG TAGCCACAGC CCTGATGAAA CTCAAAAGCC TAGGTCTTAT CCACGCTGAC 1080
 CTCAAAACAG AATACATCAT GCTGGTGGAT CCATCTAGAC AACCATACAG AGTCAAGGTC 1140
 ATCGACTTGT GTTCAGCAGC CCACGTCTCC AAGGCTGTGT GCTCCACCTA CTTGCAAGTCC 1200
 AGATATTACA GGGCCCCTGA GATCATCCTT GGTTTACCAT TTTGTGAGGC AATTGACATG 1260
 TGGTCCCTGG CTTGTGTAT TGCAGAAATT TTCTGGGTT GGCCGTTATA TCCAGGAGCT 1320
 55 TCCGAGTATG ATCAGATTCT GTATATTCCA CAAACACAGG GTTTGCCTGC TGAATATTTA 1380
 TTAAGCCCGG GGACAAAGAC AACTAGGTTT TTCAACCGTG ACACGGACTC ACCATATCCT 1440
 TTGTGGAGAC TGAAGACACC AGATGACCAT GAAGCAGAGA CAGGGATTAA GTCAAAAGAA 1500
 GCAAGAAAGT ACATTTTCAA CTGTTTAGAT GATATGGCCC AGGTGAACAT GACGACAGAT 1560
 TTGGAAGGGA GCGACATGTT GGTAGAAAAG GCTGACCGGC GGGAGTTTCA TGACCTGTTG 1620
 60 AAGAAGATGC TGACCATTGA TGTGACAAG AGAATCACTC CAATCGAAAC CCTGAACCAT 1680
 CCCTTTGTCA CCATGACACA CTTACTCGAT TTCCCCACA GCACACACGT CAAATCATGT 1740
 TTCCAGAAAC TGGAGATCTG CAAGCGTCGG GTGAATATGT ATGACACGGT GAACCAAGAGC 1800
 AAAACCCCTT TCATCACGCA CGTGGCCCCC AGCACGTCCA CCAACCTGAC CATGACCTTT 1860
 AACACACAGC TGACCATCTG CCACAACAGC GTCCCTCCT CTACCACTGC CACTATTTC 1920
 65 TTAGCCAATC CCGAAGTCTC CATACTAAAC TACCATCTA CACTTACCA GCCCTCAGCG 1980
 GCATCCATGG CTGCAGTGGC CCAGCGGAGC ATGCCCTGTC AGACAGGAAC AGCCAGATT 2040
 TGTGCCCGGC CTGACCCGTT CCAGCAAGCT CTCATCGTGT GTCCCCCGG CTTCCAAGGC 2100
 TTGCAGGCTT CTCCCTCTAA GCACGCTGGC TACTCGGTGC GAATGGAAAA TGCAGTTTCC 2160
 ATCGTCACTC AAGCCCAGG AGCTCAGCCT CTTCAGATCC AACCAAGTCT GCTTGCCAG 2220
 70 CAGGCTTGGC CAAGTGGGAC CCAGCAGATC CTGCTTCCC CAGCATGGCA GCAACTGACT 2280
 GGAGTGGCCA CCCACACATC AGTGCAGCAT GCCACCGTGA TTCCCGAGAC CATGGCAGGC 2340
 ACCCAGCAGC TGGCGGACTG GAGAAATACG CATGCTCAGC GAAGCCATTA TAATCCCATC 2400
 ATGCAGCAGC CTGCATCTAT GTGACCTTTC CAGCAGCACA GCCCTTAAAT 2460
 GTGGGTGTGG CCCACGTGAT GCGGCAGCAG CCAACCAGCA CCACCTCCTC CCGGAAGAGT 2520
 AAGCAGCACC AGTCATCTGT GAGAAATGTC TCCACCTGTG AGGTGTCTTC CTCTCAGGCC 2580
 75 ATCAGCTCCC CACCTGCTGC CAAGCGTGTG AAGGAGAACA CACCTCCCCG CTGTGCCATG 2640
 GTGCACAGTA GCGCGGCTTG CAGCACCTCG GTCACTGTGT GGTGGGGCGA CGTGGCCTCC 2700
 AGCACACCCC GCGAAGGACA GCGGCAGACA ATGTCTATT CCGACACTCC CAGCCCCACG 2760
 GTCAGCGTCA TCACCATCAG CAGTGACACG GACGAGGAGG AGGAACAGAA ACACGCCCCC 2820
 ACCAGCACTG TCTCCAAGCA AAGAAAAAAC GTCATCAGCT GTGTCAAGT CCACGACTCC 2880
 80 CCTTACTCCG ACTCCTCCAG CAACACCAGC CCTTACTCCG TGCAGCAGCG TGCTGGGCAC 2940

AACAAATGCCA ATGCCTTTGA CACCAAGGGG AGCCTGGAGA ATCACTGCAC GGGGAACCCC 3000
 CGAACCATCA TCGTGCCACC CCTGAAAACC CAGGCCAGCG AAGTATTGGT GGAGTGTGAT 3060
 AGCCTGGTGC CAGTCAACAC CAGTCACCAC TCGTCTCTCT ACAAGTCCAA GTCTCTCAGC 3120
 AACGTGACCT CCACCAGCGG TCACTCTTCA GGGAGCTCAT CTGGAGCCAT CACCTACCGG 3180
 CAGCAGCGGC CGGGCCCCCA CTTCCAGCAG CAGCAGCCAC TCAATCTCAG CCAGGCTCAG 3240
 CAGCACATCA CCACGACCG CACTGGGAGC CACCGAAGGC AGCAGGCCA CATCACTCCC 3300
 ACCATGGCCC AGGCTCCGTA TCCTTCCCG CACAACAGCC CCAGCCACGG CACTGTGCAC 3360
 CCGCATCTGG CTGCAGCCGC TGCCGCTGCC CACCTCCCCA CCCAGCCCCA CCTCTACACC 3420
 TACACTGCGC CGGCGGCCCT GGGCTCCACC GGCACCGTGG CCCACCTGGT GGCTTCGCAA 3480
 GGCTCTGCGC GCCACACCGT GCAGCACACT GCCTACCAG CCAGCATCGT CCACCAGGTC 3540
 CCCGTGAGCA TGGGCCCCCG GGTCTGCCC TCGCCACCA TCCACCCGAG TCAGTATCCA 3600
 GCCCAATTGT CCCACCAGAC CTACATCAGC GCCTCGCCAG CCTCCACCGT CTACACTGGA 3660
 TACCACATGA GCCCCGCCAA GGTCAACAG TACCCTTACA TATAAACACT GGAGGGGAGG 3720
 GAGGGAGGGA GGGGCGGAGA GAAATGGCCCG AGGGAGGAGG GAGAGAAGGA GGGAGGCGCT 3780
 CCTGGGACCG TGGGCGCTGG CCTTTTATAC TGAAGATGCC GCACACAAC AATGCAAACG 3840
 GGGCAGGGGC GGGGGGGGGG GGGGCAGAGG GCAGGGGGAC GGGTCGGGAC ACCAGTGAAA 3900
 CTTGAACCGG GAAATGGGAG GACGTAGAGC AGAGAAGAGA ACATTTTAA AAGGAAGGGA 3960
 TTAAGAGGG TGGGAAATCT ATGGTTTATA TTTTAAAAA

SEQ ID NO:282 PCI2 Protein sequence:

Protein Accession #: NP_073577

MAPVYEGMAS HVQVFSPTL QSSAFCSVKK LKVEPSSNWD MTGYGSHSKV YSQSKNIPPS 60
 QPASTIVSTS LPVNPSPSLPY EQTIVFPGST GHIVVTSASS TSVTGQVLGG PHNLMRRSTV 120
 SLLDTYQKCG LKRRKSEIEIEN TSSVQHEEH PPMIQNNASG ATVATATST ATSKNSGNS 180
 EGDYQLVQHE VLCSMTNTYE VLEFLGRGTF GQVVKCWKRQ TNEIVAIIKL KNRPSYARQG 240
 QIEVSILARL STESADDYFN VRAIECFQHK NHTCLVFEML EQNLYDFLKQ NKFSPLPLKY 300
 IRPVLQQVAT ALMKLKSGL IHADLKPENI MLVDPSRQPY RVKVIDFGSA SHVSKAVCST 360
 YLQSRYYRAP EILGLPFCE AIDMWSLGCV IAEFLGWPL YPGASEYDQI RYISQTOQLP 420
 AEYLLSAGTK TTRFFNRDITD SPYPLWRLKT PDDHEAETGI KSKEARKYIF NCLDDMAQVN 480
 MTTDLESDM LVEKADREF IDLLKKMLTI DADKRITPIE TLNHPFVTMT HLLDFPHSTH 540
 VKSCFQNMCI CKRRVNMYYDT VNQSKTPFIT HVAPSTSNL TMTFNNQLTT VHNQAPSSTS 600
 ATISLANPEV SILNYPSTLY QPSAASMAAV AQRSMPLQTG TAQICARPDP FQALIVCPP 660
 GFQGLQASPS KHAGYSVRME NAVPIVQAP GAQPLQIQPG LLAQQAUPSG TQIILLPPAW 720
 QQLTGVAHT SVQHATVIPE TMAGTQQLAD WRNTHAHGSH YNPIMQPAL LTGHVILPAA 780
 QPLNVGAHV MRQPTSTTS SRKSKQHQS VRNVSTCEVS SSQAISPPQR SKRVKENTPP 840
 RCAMVHSSPA CSTSVTCGWG DVASSTIRER QRQTIIPDT PSPTVSVITI SSDTDEEEQ 900
 KHAFTSVSK QRKNVISCVT VHDSPYSDSS SNTSPYSVQQ RAGHNNANAF DTGKSLENHC 960
 TGNPRTHVP PLKTQASEVL VECDSLVPVN TSHHSSYSKS KSSSNVTSTS GHSSGSSSGA 1020
 IYRQQRPGP HFQOQQLNL SQAQQHITD RTGSHRRQQA YIPTMAQAP YSFPHNSPSH 1080
 GTVPHLAAA AAAHLPTQP HLYTYTAPAA LGSTGTVAHL VASQGSARHT VQHTAYPASI 1140
 VHQPVSMGP RVLPSPTIHP SQYPAQFAHQ TYISASPAST VYTGYPPLSPA KVNQYPYI

SEQ ID NO:283 PB1 DNA SEQUENCE

Nucleic Acid Accession#: NM_017700

Coding sequence: 147-806 (underlined sequence corresponds to start and stop codon)

1 11 21 31 41 51
 AGTCACAGCC AGGTAACCTT GGAGTGAAGC GGTTTAGTTA GAAGGGAGCA GATAAACTCG 60
 TCACTCTAGT AGCTTTAACC CTCACCTGTA GGCACCTTAG CAATCAGCCA TTGCCTGCAA 120
 GCCTCCAAAG CTTGCTCTTT CCTAATATGG AGCCCAAAGA AGCCACTGGG AAAGAAAACA 180
 TGGTCACCAA GAAAAGAAAT CTGGCCTTCT TGAGGTCTAG ACTCTATATG CTGGAGAGAA 240
 GGAAGACTGA CACTGTGGTT GAGAGCAGTG TTTCTGGGGA CCACTCTGGC ACCTTGAGGA 300
 GGAGCCAAATC TGACAGGACC GAATACAACC AGAAATTACA AGAAAAGATG ACTCCACAGG 360
 GTGAGTGTTC TGTAGCTGAG ACCTTAACCC CAGAGGAAGA GCATCATATG AAGAGGATGA 420
 TGGCAAAGCG GGAAGAGATG ATTAAGGAGC TGATACAGAC AGAAAAGGAT TATCTCAATG 480
 ATCTAGAGCT GTGTGTTAGG GAAGTGGTTC AGCCCTGAG AAATAAAAAG ACTGATAGGC 540
 TGGATGTGGA TAGCTTGTTT AGCAACATTG AGTCCGTGCA TCAGATATCA GCCAAGCTGC 600
 TGTCAATGTT GGAAGAGGCC ACAACAGACG TGGAAACCGG CATGCAAGTA ATGGGAGAAG 660
 TATTCTTGCA GATTAAAGGG CCACTGGAAG ATATTATATA AATCTACTGC TATCACCATG 720
 ATGAAGCACA TAGTACTACTG GAGTCTATG AAAAGGAAGA AGAGCTGAAG GAACATTTGA 780
 GCCACTGTAT CCAGTCCCTTA AAGTAAGGCC TTTTCAATG ATGATTCCTA TCTCCTCTCA 840
 GTTGCTTAGC AGGGAACATT TTAATGGGAT GTAGATGAAA GGTCTCACAT AAATCCTATG 900
 TTTTATGAGA CTTGCTGGGA GCTCTGCTTT GCATTCCTTT TATAAAAAGC TGACATGCCA 960
 GAAGCCCTGA TTGACTTTT TTCCCTCTGC GAGAATGACT AAAAATAACA TGGAAAGAAGA 1020
 TTTAGAGCTC TGCAGCGATT GAAAATGCA ATATCAAAAT ATAAATGTG GAAGAAAAGC 1080
 CTCCTCTTAA AGCTATTGTA ACTTGCCCTGG CCCACGCTAG TTCAAGGATT ATGTGAGATA 1140
 ACACGTGGCC CCATGACCAC TGGAGCACAT GGGTTAATGG AGTTAGGGGA ATGGCCTACA 1200
 ACTCTGCATG GCCGTCTTCT TTCCCCAAAC TCACTGTGGG GAGATGGGGT AAGACAAGTC 1260
 AGGCCTTGTG AAAGTTAGTT TCAGAACAAAT TACTCATGCC TTCTTTCTC ATCCCTAAAA 1320
 CATTTGGTGGG GGAGCTACAC AATGTACTTT TTCTTTTCTA GAGGAAGTAT CTATTCACTG 1380
 TGAAAACTCG AAAAATATAA CAAAGTATGT GTAAGATAAA AACCCCTTGC TATTTCAAAA 1440
 AAAAAAAAAA AAAAAAAAAA AAAA

SEQ ID NO:284 PB1 Protein sequence:

Protein Accession #: NP_060170

1 11 21 31 41 51
 420

	MEPKEATGKE	NMVTKKKNLA	FLRSRLYMLE	RRKTDTVVES	SVSGDHSGLT	RRSQSDRTEY
	NQKLQEKMTPT	QGECSVAETL	TPREEHHMKR	MMAKREKIHK	ELIQTEKDYL	NDLELCVREV
5	VQPLRNKKT	RLDVSLSFSN	IESVHQISAK	LLSLLEBAT	DVEPAMQVIG	EVFLQIKGFL
	EDIIYKIYCYH	HDEAHSILES	YEKEEELKEH	LSHCILQSLK		

SEQ ID NO:285 PBQ9 DNA SEQUENCE

Nucleic Acid Accession#: X66534
Coding sequence: 523-2676 (underlined sequence corresponds to start and stop codon)

	1	11	21	31	41	51	
15	CCCTTATGGC	GATTGGGCGG	CTGCAGAGAC	CAGGACTCAG	TTCCCTTGCC	CTAGTCTGAG	60
	CCTAGTGGGT	GGGACTCAGC	TCAGAGTCAG	TTTTCAGAAG	CAGGTTCAG	TTGCAGAGTT	120
	TTCTTACACT	TTTCCGCGC	TAGAGCAGCG	AGCAGCCTGG	AACAGACCCA	GGCGGAGGAC	180
	ACCTGTGGGG	GAGGGAGCGC	CTGGAGGAGC	TTAGAGACCC	CAGCCGGGCG	TGATCTCACC	240
	ATGTGCGGAT	TTGCGAGGCG	CGCCCTGGAG	CTGCTAGAGA	TCCGAAGCA	CAGCCCCGAG	300
20	GTGTGCGAAG	CCACCAAGAC	TGCGGCTCTT	GGAGAAAGCG	TGAGCAGGGG	GCCACCGCGG	360
	TCCTCCGGCT	GTCTGCACCC	TGTCGCCTGA	GCTGCCTGAC	AGTGACAAATG	ACATCCCACT	420
	TACCAGTGTC	CTTGAATTGA	TAGTGGCTTC	TGTTTGTCAG	TCTCATATAA	GAACCTACAGC	480
	TCATCAGGAG	GAGATCGCAG	CAGGGTAAGA	GACACCAACA	CCATGTTCTG	CACGAAGCTC	540
	AAGGATCTCA	AGATCACAGG	AGAGTGCTCT	TTCTCCCTTAC	TGGCACCAGG	TCAAGTTCCCT	600
25	AACGAGTCTT	CAGAGGAGGC	AGCAGGAAGC	TCAGAGAGCT	GCAAAGCAAC	CGTGCCCATC	660
	TGTCAAGACA	TTCTGAGAA	GAACATACAA	GAAAGTCTTC	CTCAAAGAAA	AACGAGTCGG	720
	AGCCGAGTCT	ATCTTCACAC	TTTGGCAGAG	AGTATTGCA	AACCTGATT	CCCAGAGTTT	780
	GAACGGCTGA	ATGTTGCAC	TCAGAGAAAC	TTGGCAAGC	ACAAATATAA	AGAAAGCAGG	840
	AAATCTTTTG	AAAGAGAAGA	CTTTGAAAAA	ACAATGTCAG	AGCAAGCAGT	GCAGCAGAGT	900
30	CCAGTGGAGT	TATCAAGAA	TCTCTTGGTG	AAGAGGTTT	TAAATATATG	TACGAGGAAG	960
	ATGAAAACAT	CTTTGGGGTG	GTTGGAGGCA	CCCTTAAAGA	TTTTTAAACA	GCTTCAGTAC	1020
	CCCTCTGAAA	CAGAGCAGCC	ATTGCCAAGA	AGCAGGAAAA	AGGGGCGAGT	TGAGGACGCC	1080
	TCCATCTTAT	GCTTGATATA	GGAGGATGAT	TTCTTACATG	TTTACTACTT	CTTCCCTAAG	1140
	AGAACCACCT	CCCTGATCTT	TCCCGGCATC	ATAAAGGCAG	CTGCTCAGCT	ATTATATGAA	1200
35	ACGGAAGTGG	AAGTGTCTGT	AATGCCCTCC	TGCTTCCATA	ATGATTGCA	CGAGTTTGTG	1260
	AATCAGCCCT	ACTTGTGTGA	CTCCGTTTAC	ATGAAAAGCA	CCAAGCCATC	CCTGTCCCCC	1320
	AGCAAAACCC	AGTCCCTGCT	GGTGATTTCC	ACATCGCTAT	TCTGCAAGAC	ATTTCATATC	1380
	CATTTCATGT	TTGACAAGA	TATGACAAAT	CTGCAATTTG	GCAATGGCAT	CAGAAGGCTG	1440
40	ATGAACAGGA	GAGACTTTCA	AGGAAAGCCT	AATTTTGAAT	ACTTTGAAAT	TCTGACTCCA	1500
	AAAAATCAAC	AGACCTTTAG	CGGGATCATG	ACTATGTTGA	ATATGCAGTT	TGTTGTACGA	1560
	GTGAGGAGAT	GGGACAACAT	TGTGAAGAAA	TCTTCAAGGG	TTATGGACCT	CAAAGGCCAA	1620
	ATGATCTACA	TTGTTGAATC	CAGTGCAATC	TTGTTTGTGG	GGTCACCCCT	TGTGGACAGA	1680
	TTAGAAGATT	TTACAGGACG	AGGGCTCTAC	CTCTCAGACA	TCCCAATTCA	CAATGCACTG	1740
45	AGGGATGTGG	TCCTAATAGG	GGAACAAGCC	CGAGCTCAAG	ATGGCCTGAA	GAAGAGGCTG	1800
	GGGAAGCTGA	AGGCTACCTT	TGAGCAAGCC	CACCAAGCCC	TGGAGGAGGA	GAAGAAAAAG	1860
	ACAGTAGAGT	TTCTGTGCTC	CATATTTCCC	TGTGAGGTTG	CTCAGCAGCT	GTGGCAAGGG	1920
	CAAGTTGTGC	AAGCCAAAGA	GTTCAAGTAT	GTCACCATGC	TCTTCTCAGA	CATCGTTGGG	1980
	TTCACTGCCA	TCTGCTCCCA	GTGCTCACCG	CTGCAGGTCA	TCACCATGCT	CAATGCACTG	2040
50	TACACTCGCT	TCGACACGCA	GTGTGGAGAG	CTGGATGTCT	ACAAGGTGGA	GACCATTTGCG	2100
	ATGCCTATTG	TGTGGCTTGG	GGGATTACAC	AAAGAGAGTG	ATACTCATGC	TGTTTCAGATA	2160
	GCCTGATGAG	CCCTGAAGAT	GATGGAGCTC	TCTGATGAAG	TTATGTCTCC	CCATGGAGAA	2220
	CCTATCAAGA	TGCGAATTGG	ACTGCACTCT	GGATCAGTTT	TTGCTGGCGT	CGTTGGAGTT	2280
	AAAAATGCCCC	GTTACTGTCT	TTTTGGAAAC	AATGTCACCT	TGGCTAACAA	ATTTGAGTCC	2340
55	TGCAGTGTAC	CAGGAAAAAT	CAATGTCAGC	CCAACAACCT	ACAGATTACT	CAAAGACTGT	2400
	CTTGTGTTTCG	TGTTTACCCC	TCGATCAAGG	GAGGAACCTC	CACCAAACTT	CCCTAGTGAA	2460
	ATCCCCGGAA	TCTGCCATTT	TCTGGATGCT	TACCAACAAG	GAACAAACTC	AAAACCATGC	2520
	TTCCAAAAGA	AAGATGTGGA	AGATGCAAGC	CAATTTTTTA	GGCAAAGCAT	CAGGAATAGA	2580
	TTAGCAACCT	ATATACCTAT	TTATAAGTCT	TTGGGGTTTG	ACTCATTGAA	GATGTGTAGA	2640
60	GCCTCTGAAA	GCACCTTTAGG	GATTGTAGAT	GGCTAACCAAG	CAGTATTAAA	ATTTTCAAGG	2700
	CCAAGTCACA	ATCTTCTTCC	TGTTTAAACAT	GACAAAAATG	ACTCACTTCA	GTAATTCAGC	2760
	TCTTCAAGAA	AAAAAATAAA	ACCTTAAAAA	GCTACTTTTG	TGGGAGTATT	TCTATTATAT	2820
	AACCAGCACT	TACTACCTGT	ACTCAAAAT	CAGCACCTTG	TACATATATC	AGATAATTGT	2880
	AGTCAATTTG	ACAAACTGAT	GGAGTCACCT	GCAATCTCAT	ATCCTGGTGG	AATGCCATGG	2940
65	TTATTAAAGT	GTGTTTGTGA	TAGTTGTCTG	CAAAAAATAA	AAAAAATAA	AAAAAATAA	3000
	AAAA						

SEQ ID NO:286 PBQ9 Protein sequence:

Protein Accession #: Q02108

	1	11	21	31	41	51	
70	MFCTKLKDLK	ITGCEPFSLL	APGQVFNES	EBAAGSSESC	KATVPICQDI	PEKNIQESLP	60
	QRKTSRSRVY	LHTLAESICK	LIFPEFERLN	VALQRTLAKH	KIKESRKSLE	REDFEKTIAE	120
	QVAAGVPEVE	VIKESLGEVE	FKICYEEDEN	ILGVVGGTLK	DFLNSFSTLL	KQSSHCQEAG	180
75	KRGRLDASI	LCLEDKEDFL	HVYFFPKRT	TSLLILPGIHK	AAAHVLYETE	VEVSLMPPCF	240
	HNDCSEFVNO	PYLLYSVHMK	STKPSLSPSK	QSSSLVIPTS	LFCKTFPFHF	MFDDKMTILQ	300
	FGNGIRRLMN	RRDFQGGPNF	EEYFEILTPK	INQTFSGIMT	MLNMQFVVRV	RRWDNSVKKK	360
	SRVMDLKGQM	IYIVESSAIL	FLGSPCVDR	EDFTGRGLYL	SDIPIHNAIR	DVVILIGEQR	420
	AQDGLKKRLG	KLKATLEQAH	QALEEEKKKT	VDLLCSIFPC	EVAQQLWQQG	VVQAKKFSNV	480
80	TMLFSDIVGF	TAICSQCSPL	QVITMLNALY	TRFDQQCCEL	DVYKVTIGD	AYCVAGGLHK	540

ESDTHAVQIA LMLAKMMELS DEVMSPHGEP IKMRIGLHSG SVFAGVVGVK MPRYCLFGNN 600
 VTLANKFESC SVPRKINVSF TTYRLKDCP GFVFTPRSRE ELPPNFPSEI PGICHFLDAY 660
 QQGTNSKPCF QKKDVEDGNA NFLGKASGID

5

SEQ ID NO:287 PFD2 DNA SEQUENCE

Nucleic Acid Accession#: NM_000720
 Coding sequence: 119-6664 (underlined sequence corresponds to start and stop codon)

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

1	11	21	31	41	51	
AGAATAAGGG	CAGGGACCGC	GGCTCCTATC	TCTTGGTGAT	CCCCTTCCCC	ATTCCGCCCC	60
CGCCTCAACG	CCCAGCACAG	TGCCCTGCAC	ACAGTAGTCG	CTCAATAAAT	GTTCGTGGAT	120
GATGATGATG	ATGATGATGA	AAAAAATGCA	GCATCAACGG	CAGCAGCAAG	CGGACCACGC	180
GAACGAGGCA	AACTATGCAA	GAGGCACCA	ACTTCCTCTT	TCTGGTGAAG	GACCAACTTC	240
TCAGCCGAAT	AGCTCCAAAG	AAACTGTCTT	GTCTTGGCAA	GCTGCAATCG	ATGCTGCTAG	300
ACAGGCCAAG	GCTGCCCAAA	CTATGAGCAC	CTCTGCACCC	CCACCTGTAG	GATCTCTCTC	360
CCAAAGAAAA	CTCAGCAAT	ACGCCAAGAG	CAAAAAACAG	GGTAACTCGT	CCAACAGCCG	420
ACCTGCCCGC	GCCCTTTTCT	GTTTATCACT	CAATAACCCC	ATCCGAAGAG	CCTGCATTAG	480
TATAGTGGAA	TGGAACCAT	TTGACATATT	TATATTATTG	GCTATTTTTC	CCAATTGTGT	540
GGCCTTAGCT	ATTTACATCC	CATTCCCTGA	AGATGATTCT	AATTCAACAA	ATCATAACTT	600
GGAAAAAGTA	GAATATGCGT	TCCTGATTAT	TTTTACAGTC	GAGACATTTT	TGAAGATTAT	660
AGCGTATGGA	TTATTGCTAC	ATCCTAATGC	TTATGTTAGG	AATGGATGGA	ATTTACTGGA	720
TTTTGTATTA	GTAATAGTAG	GATTGTTTAG	TGTAATTTTC	GAACAATTAA	CCAAAGAAAC	780
AGAAGGCGGG	AACCACTCAA	GCGGCAAAAT	TGGAGGCTTT	GATGTCAAAG	CCCTCCGTGC	840
CTTTCGAGTG	TTGCGACCA	TTGCACTAGT	GTGAGGGGTG	CCCAGTTTAC	AAGTTGTCCT	900
GAATCCATT	ATAAAAGCCA	TGGTTCCTCT	CCTTCACATA	GCCCTTTTGG	TATTTATTGT	960
AATCATAATC	TATGCTATTA	TAGGATGGA	ACTTTTATAT	GGAAAAATGC	ACAAAAATGC	1020
TTTTTTTGGT	GACTCAGATA	TCGTAGCTGA	AGAGGACCCA	GCTCCATGTG	CGTTCTCAGG	1080
GAATGGACCC	CAGTGTACTG	CCAATGGCAC	GGAAATGTAG	AGTGGCTGGG	TTGGCCCCGA	1140
CGGAGGCATC	ACCAACTTTG	ATAACTTTGC	CTTTGCCATG	CTTACTGTGT	TTCACTGCAT	1200
CACCATGGAG	GCTTGGACAG	ACGTGCTCTA	CTGGGTAAAT	GATGCGATAG	GATGGGAATG	1260
GCCATGGGTG	TATTTTGTTA	GTCTGATCAT	CCTTGGCTCA	TTTTTCGTCC	TTAACCTGGT	1320
TCTTGGTGTC	CTTAGTGGAG	AATTCTCAA	GGAAAGAGAG	AAGGCAAAAG	CACGGGGAGA	1380
TTTCCAGAAG	CTCCGGGAGA	AGCAGCAGCT	GGAGGAGGAT	CTAAAGGGCT	ACTTGGATTG	1440
GATCACCCAA	GCTGAGGACA	TCGATCCGGA	GAATGAGGAA	GAAGGAGGAG	AGGAAGGCCA	1500
ACGAATAACT	AGCATGCCCA	CCAGCGAGAC	TGAGTCTGTG	AACACAGAGA	ACGTCAGCGG	1560
TGAAGGCGAG	AACCGAGGCT	GCTGTGGAAG	TCTCTGGTGC	TGGTGGAGAC	GGAGAGGCGC	1620
GGCCAAGCGG	GGGCCCTCTG	GGTGTGCGCG	GTGGGGTCAA	GCCATCTCAA	AATCCAAACT	1680
CAGCCGACGC	TGGCGTCGCT	GGAACCGATT	CAATCGCAGA	AGATGTAGGG	CCGCCGTGAA	1740
GTCTGTACAG	TTTTACTTGC	TGGTATTCGT	CCTGGTGTTC	CTGAACACCT	TAAACCTTTC	1800
CTCTGAGCAC	TACAAATCAG	CAGATGGTTC	GACACAGATT	CAAGATATTG	CCAACAAAGT	1860
CCTCTTGGCT	CTGTTCACCT	GCGAGATGCT	GGTAAAAATG	TACAGCTTGG	GCCTCCAAAG	1920
ATATTTCGTC	TCTCTTTTCA	ACCGGTTTGA	TTGCTTTCGT	GTGTGTGGTG	GAATCACTGA	1980
GACGATCCTG	TGGAACTGG	AAATCATGTC	TCCCTTGGGG	ATCTCTGTGT	TTCCGTTGTT	2040
GGCCTCTTCA	AGAATCTTCA	AAGTGACCAG	GCACCTGGAC	TCCCTGAGCA	ACTTAGTGGC	2100
ATCCTTATTA	AACTCCATGA	AGTCCATCGC	TTGCTGTGTC	CTTCTGCTTT	TTCTCTTCAT	2160
TATCATCTTT	TCCTTGCTTG	GGATGCAGCT	GTTTGGCGGC	AAGTTTAATT	TTGATGAAAC	2220
GCAAAACCAAG	CGGAGCACCT	TTGACAATTT	CCCTCAAGCA	CTTCTCAGAG	TGTTCCAGAT	2280
CCTGACAGGC	GAAGACTGGA	ATGCTGTGAT	GTACGATGGC	ATCATGGCTT	ACGGGGGCCC	2340
ATCCTCTTCA	GGAATGATCG	TCTGCATCTA	CTTCATCATC	CTCTTCATTT	GTGGTAACTA	2400
TATTCTACTG	AATGTCTTCT	TGGCCATCGC	TGTAGACAAT	TTGGCTGATG	CTGAAAGTCT	2460
GAACACTGCT	CAGAAAGAAG	AAGCGGAAGA	AAAGGAGAGG	AAAAAGATTG	CCAGAAAAGA	2520
GAGCCTAGAA	AATAAAAAGA	ACAACAACCC	AGAAGTCAAC	CAGATAGCCA	ACAGTGACAA	2580
CAAGGTTACA	ATTGATGACT	ATAGAGAAGA	GGATGAAGAC	AAGGACCCCT	ATCCGCCCTT	2640
CGATGTGCCA	GTAGGGGAAG	AGGAAGAGGA	AGAGGAGGAG	GATGAACCTG	AGGTTCTCTG	2700
CGGACCCCGT	CCTCGAAGGA	TCTCGGAGTT	GAACATGAAG	GAAAAAATTG	CCCCCATCCC	2760
TGAAGGGAGC	GCTTTCTTCA	TTCTTAGCAA	GACCAACCCG	ATCCGCGTAG	GCTGCCACAA	2820
GCTCATCAAC	CACCAACATC	TCACCAACCT	CATCCTTGTC	TTTCATCATG	TGAGCAGCGC	2880
TGCCCTTGGC	GAGAGGACC	CCATCCGCAG	CCACTCCTTC	CGGAACACGA	TACTGGGTTA	2940
CTTTGACTAT	GCCTTCACAG	CCATCTTTAC	TGTTGAGATC	CTGTTGAAGA	TGACAACTTT	3000
TGGAGCTTTT	CTCCACAAGG	GGGCTTCTG	CAGGAACCTC	TTCAATTTCG	TGGATATGCT	3060
GGTGGTTGGG	GTGCTCTTGG	TGTCATTTGG	GATTCAATCC	AGTGCCATCT	CCGTTGTGAA	3120
GATTCTGAGG	GTCTTAAGGG	TCCTGCGTCC	CCTCAGGGCC	ATCAACAGAG	CAAAAGGACT	3180
TAAGCACGTG	GTCCAGTGGC	TCTTCGTGGC	CATCCGGACC	ATCGGCAACA	TCATGATCGT	3240
CCTACCCCTC	CTGACGTTCA	TGTTTGGCTG	TATCGGGGTC	CAGTTGTTTC	AGGGGAAGTT	3300
CTATCGCTGT	ACGGATGAAG	CCAAAAGTAA	CCCTGAAGAA	TGCAGGGGAC	TTTTTCATCT	3360
CTACAAGGAT	GAGGATGTTT	ACAGTCTCTG	GGTCCGTGAA	CGGATCTGGC	AAAACAGTGA	3420
TTTCAACTTC	GACAACGTCC	TCTCTGCTAT	GATGGCGCTC	TTACAGTCTT	CCAGTTTGA	3480
GGGCTGGCCT	GCGTTGCTGT	ATAAAGCCAT	CGACTCGAAT	GGAGAGAACA	TGGGCCCAAT	3540
CTACAACCAC	CGCGTAGGAG	TCTCCATCTT	CTTCATCATC	TACATCATCA	TTGTAGCTTT	3600
CTTCAATGATG	AACATCTTTT	TGGGCTTTGT	CATCGTTACA	TTTCAGGAAC	AAGGAGAAAA	3660
AGAGTATAAG	AACCTGTAGC	TGGACAAAAA	TCAGCGTCAG	TGTGTTGAAT	ACGCCTTGAA	3720
AGCACGTCCC	TTGCGGAGAT	ACATCCCCAA	AAACCCCTAC	CAGTACAAGT	TCTGGTACGT	3780
GGTGAACCTC	TGCGCTTTTC	AATACATGAT	GTTTGTCTTC	ATCATGCTCA	ACACACTCTG	3840
CTTGGCCCATG	CAGACGTGAA	AGCAGTCCAA	GATGTTCAAT	GATGCCATGG	ACATCTGAA	3900
CATGGTCTTC	ACCGGGGTGT	TCACCGTCGA	GATGGTTTTC	AAAGTCATCG	CATTTAAGCC	3960
TAAGGGGTAT	TTTAGTGACG	CCTGGAACAC	GTTTGACTCC	CTCATCGTAA	TGGCAGCAT	4020
TATAGACGTG	CGCCCTACGC	AAGCGGACCC	AACTGAAAGT	GAAAAATGTC	CTGTCCCAAC	4080

5 TGCTACACCT GGGAACTCTG AAGAGAGCAA TAGAATCTCC ATCACCTTTT TCCGCTCTTT 4140
 CCGAGTGTAT CGATTGGTGA AGCTTCTCAG CAGGGGGGAA GGCATCCGGA CATTGCTGTG 4200
 GACTTTTATT AAGTCCTTTC AGGCGCTCCC GTATGTGGCC CTCCTCATAG CCATGCTGTT 4260
 CTTCATCTAT GCGGTCAATTG GCATGCAGAT GTTTGGGAAA GTTGCCATGA GAGATAACAA 4320
 CCAGATCAAT AGGAACAATA ACTTCCAGAC GTTCCCCAG GCGGTGCTGC TGCTCTTCAG 4380
 GTGTGCAACA GGTGAGGCTT GGCAGGAGAT CATGCTGGCC TGCTCCCGAG GGAAGCTCTG 4440
 TGACCTTGAG TCAGATTACA ACCCCGGGGA GGAGTATACA TGTGGGAGCA ACTTTGCCAT 4500
 10 TGCTATTTTC ATCAGTTTTT ACATGCTCTG TGCAATTTCTG ATCATCAATC TGTTTGTGGC 4560
 TGTCATCATG GATAATTTTC ACTATCTGAC CCGGGACTGG TCTATTTTGG GGCCTCACCA 4620
 TTTAGATGAA TTCAAAGAA TATGGTCTAGA ATATGACCTT GAGGCAAGG GAAGGATAAA 4680
 ACACCTTGAT GTGGTCACTC TGCTTCGACG CATCCAGCCT CCCCTGGGGT TTGGGAAGTT 4740
 ATGTCCACAC AGGGTAGCGT GCAAGAGATT AGTTGCCATG AACATGCCCT TCAACAGTGA 4800
 CGGGACAGTC ATGTTTAATG CAACCTGTGT TGCTTTGGTT CGAACGGCTC TTAAGATCAA 4860
 15 GACCGAAGGG AACCTGGAGC AAGCTAATGA AGAACTTCGG GCTGTGATAA AGAAAATTTG 4920
 GAAGAAAACC AGCATGAAAT TACTTGACCA AGTTGTCCCT CCAGCTGGTG ATGATGAGGT 4980
 AACCGTGGGG AAGTTCTATG CCACTTTCTT GATACAGGAC TACTTTAGGA AATTCAAGAA 5040
 ACGGAAAAGAA CAAGGACTGG TGGGAAAGTA CCCTGCGAAG AACACCACAA TTGCCCTACA 5100
 GGGGGGATTA AGGACACTGC ATGACATTGG GCCAGAAATC CGGCGTGCTA TATCGTGTGA 5160
 20 TTTGCAAGAT GACGAGCTTG AGGAAACAAA ACGAGAAGAA GAAGATGATG TGTTCAAAAG 5220
 AAATGGTGCC CTGCTTGGAA ACCATGTCAA TCATGTTAAT AGTGATAGGA GAGATTCCCT 5280
 TCAGCAGACC AATACCACCC ACCGTCCCCT GCATGTCCAA AGGCCCTCAA TTCCACCTGC 5340
 AAGTGATACT GAGAAACCCG TGTTCCTCC AGCAGGAAAT TCGGTGTGTC ATAACCATCA 5400
 TAACCATTAAT TCCATAGGAA AGCAAGTTCC CACCTCAACA AATGCCAATC TCAATAATGC 5460
 25 CAATATGTCC AAAGCTGCCC ATGGAAGCG GCCCAGCATT GGGAACTTTG AGCATGTGTC 5520
 TGAAAATGGG CATCATCTCT CCCACAAGCA TGACCGGGAG CCTCAGAGAA GGTCCAGTGT 5580
 GAAAGAAACC CGCTATTTAT AACTTTACAT TAGGTCCGAC TCAGGAGATG AACAGCTCCC 5640
 AACTATTTCG CGGGAAGACC CAGAGATACA TGGCTATTTT AGGGACCCCT ACTGCTTGGG 5700
 GGAGCAGGAG TATTTAGTA GTGAGGAATG CTACGAGGAT GACAGCTCGC CCACCTGGAG 5760
 CAGGCAAAAC TATGGCTACT ACAGCAGATA CCCAGGCAGA AACATCGACT CTGAGAGGCC 5820
 30 CCGAGGCTAC CATCATCCCC AAGGATTCTT GGAGGACGAT GACTCGCCCG TTTGCTATGA 5880
 TTCACGGAGA TCTCCAAGGA GACGCCCTACT ACCTCCCAAC CCAGCATCCC ACCGGAGATC 5940
 CTCCTTCAAC TTTGAGTGGC TGCGCCGGCA GAGCAGCCAG GAAGAGGTCC CGTCTGTCTC 6000
 CATCTTCCCC CATCGCACGG CCTGTCTCTT GCATCTAATG CAGCAACAGA TCATGGCAGT 6060
 35 TGCCGGGCTTA GATTCAAGTA AAGCCCAAGAA GTACTCACCG AGTCACTCGA CCCGTCGTG 6120
 GGGCACCCCT CATCAACACC CTCCCACCG GGACTGGACA CCGTGTACA CCCCTTGAT 6180
 CCAAGTGGAG CAGTCAGAGG CCCTGGACCA GGTGAACGGC AGCCTGCCGT CCCTGCACCG 6240
 CAGCTCTCTG TACACAGACG AGCCCGACAT CTCCTACCGG ACTTTACAC CAGCCAGCCT 6300
 GACTGTCCCC AGCAGCTTCC GGAACAAAAA CAGCGACAAG CAGAGGAGTG CGGACAGCTT 6360
 40 GGTGGAGGCA GCTCTGATAT CCGAAGGCTT GGGACGCTAT GCAAGGGACC CAAAATTTGT 6420
 GTACAGCAACA AAACACGAAA TCGCTGATGC CTGTGACCTC ACCATCGACG AGATGGAGAG 6480
 TGACGCCAGC ACCCTGCTTA ATGGGAACGT GCGTCCCGGA GCCAACGGGG ATGTGGGCCC 6540
 CCTCTCACAC CGGCAGGACT ATGAGCTACA GGACTTTGGT CCTGGCTACA GCGACGAAGA 6600
 GCCAGACCTT GGGAGGAGT AGGAGGACCT GCGGGATGAA ATGATATGCA TCACCACCTT 6660
 45 GTAGCCCCCA GCGAGGGGCA GACTGGCTCT GGCCTCAGGT GGGGCGCAGG AGAGCCAGGG 6720
 GAAAGTGGCC TCATAGTTAG GAAAGTTTAT GCACTAGTTG GGAGTAATAT TCAATTAAT 6780
 AGACTTTTGT ATAGAGATG TCATGCTCTA AGAAAGCCAT AAACCTGGTA GGAACAGGTC 6840
 CCAAGCGGTT GAGCCTGGCA GAGTACCATG CGCTCGGCCC CAGCTGCAGG AAACAGCAGG 6900
 50 CCGCGCCCTC TCACAGAGGA TGGGTGAGGA GGCCAGACCT GCCCTGCCCT ATTGTCCAGA 6960
 TGGGCACTGC TGTGGAGTCT GCTTCTCCCA TGTACCAGGG CACCAGGCCC ACCCAACTGA 7020
 AGGCATGGCG GCGGGGTGCA GGGGAAAGTT AAAGGTGATG ACGATCATCA CACCTCGTGT 7080
 CGTTACTCTA GCCATCGGTG TAGCATATCA GTCACCTGGC CCAACATATC CATTTTTAAA 7140
 CCTTTTCCCC CAATACACT GCGTCTTGGT TCCTGTTTAG CTGTTCTGAA ATA

SEQ ID NO:288 PFD2 Protein sequence:

Protein Accession #: A38198

55 1 11 21 31 41 51
 | | | | |
 60 MMMMMMKM QHQRQQQADH ANEANYARGT RLPLSGEGPT SQPNSSKQTV LSWQAAIDAA 60
 RQAKAAQTMS TSAPPPVGS L SQRKRQYAK SKKQGNSSNS RPARALFCLS LMNPIRRACI 120
 SIVEWKPFDI FILLAI FANC VALAIYIPFP EDDSNSTNHN LEKVEYAF LI IFTVETFLKI 180
 IAYGLLLHPN AYVRNGWNLL DFVIVIVGLF SVILEQLTKE TEGGNHSSGK SGGFDVKALR 240
 AFRVLRPLRL VSGVPSLQV LNSIIKAMVP LLHIALLVLF VIIIIYAIIGL ELFIGKMHKT 300
 65 CFFADSDIVA EEDFAPCAFS GNGRQCTANG TECSRGWVGP NGGITNFDNF AFAMLTVPFQC 360
 ITMEGWTDLV YWVNDIAGWE WPVVYFVSLI ILGSFFVLNL VLGVLSGEFS KEREKAKARG 420
 DFQKLRKQCG LEEDLKG YLD WITQAEIDID ENEEBEGGEEG KRNTSMPTSE TESVNTENV S 480
 GEGENRGCCG SLWCWRRRRG AAKAGPSGCR RWGQAISKSK LSRRRWRWRNR FNRRRCRAAV 540
 KSVTFYWLVI VLVFLNLT LI SSEHYNQPDW LTQIQDIANK VLLALFTCEM LVKMYSLGLQ 600
 AYFVSLFNRF DCFVVCGGIT ETILVELEIM SPLGISVFR C VRLLRIFKVT RHWTSLSNLV 660
 70 ASLLNSMKSI ASLLLLLFLF I IIFSLLMQ LFGGKPNFDE TQTKRSTFDN FPQALLTVFQ 720
 ILTGEDWNAV MYDGIIMYGG PSSSGMIVCI YFII LFCGN YILLNVFLAI AVDNLADAES 780
 LNTAQKEEAE EKERKKIARK ESLENKKNK PEVNQIANS D NKVTIDDYRE EDEKDFYPP 840
 CDVPVGEEEE EEEDEPEV AGPRPRISE LNMKEKIAP I PEGSAFFILS KTNPIRVGCH 900
 75 KLINHHIFTN LILVFLMLSS AALAAEDPIR SHSFRNTILG YFDYAF TAIF TVEILLKMTT 960
 FGAFLLHKGAF CRNVNLLDM LVGVSLVSF GIQSSAISV KILRLVLRVL PLRAINRAKG 1020
 LKHVVQCQFV AIRTIGNIMI VTTLLQFMFA CIGVQLFKGK FYRCTDEAKS NPEECRGLFI 1080
 LYKGDVDVSP VVRERIWNQS DFNFDNLVLSA MMALFTVSTF EGWPALLYKA IDSNGENIGP 1140
 IYNHRVEISI FFIIYI IIVA FFMNIFVGF VIVTFQEQGE KEYKNCELDK NQRQCV EYAL 1200
 80 KARPLRRYIP KNPYQYKFWY VVNSSPF EYM MFVLIMLNTL CLAMQHYEQS KMFNDAMDIL 1260
 NMVFTGVFTV EMVLKVIAFK PKGYFS DAWN TFDSLIVIGS IIDVALSEAD PTESENVFVP 1320

5	TATPGNSEES	NRISITFFRL	FRVMRLVKLL	SRGEGIRTL	WTFIKSFQAL	PYVALLIAML	1380
	FFIYAVIGMQ	MFGKVMARDN	NQINRNNNFQ	TFFQAVLLLF	RCATGEAWQE	IMLACLPGLK	1440
	CDPESDYNPG	EEYTCGSNFA	IVYFISFYML	CAFLIINLFV	AVIMDNFDYL	TRDWSILGPH	1500
	HLDEFKRIWS	EYDPEAKGRI	KHLDVVTLRL	RIQPPLGFGK	LCPHRVACKR	LVMNMPLNS	1560
	DGTVMFNATL	FALVVRTALKI	KTEGNLEQAN	EELRAVIKKI	WKKTSMKLLD	QVVPAGDDE	1620
	VTVGKIFYATF	LIQDYFRKFK	KRKEQGLVGK	YPAKNTTIAL	QAGLRTLHDI	GPEIRRAISC	1680
	DLQDEPEET	KREEEDDVFK	RNGALLGNHV	NHVNSDRRDS	LQQTNTTHRP	LHVQRPSIPP	1740
	ASDTEKPLFP	PAGNSVCHNH	HNHNSIGKQV	PTSTNANLNN	ANMSKAAHGK	RPSIGNLEHV	1800
10	SENGHSSHK	HDREPORRSS	VKRTRYET	IRSDSGDEQL	PTICREDPEI	HGYFRDPHCL	1860
	GEQEFSSSE	CYEDDSPTW	SRQNYGYYSR	YFGRNIDSER	PRGYHHFQGF	LEDDDSFVCY	1920
	DSRRSPRRRL	LPPTPASHRR	SSFNFECLRR	QSSQEEVPSS	PIFPHRTALP	LHLMQQQIMA	1980
	VAGLDSSKAQ	KYSPSHSTRS	WATPPATPPY	RDWTPCYTPL	IQVEQSEALD	QVNGSLPSLH	2040
	RSSWYTDEPD	ISYRTFTPAS	LTVPSSFRNK	NSDKQRSADS	LVEAVLISEG	LGRYARDPKF	2100
15	VSATKHEIAD	ACDLTIDEME	SAASTLLNGN	VRPRANGDVG	PLSHRQDYEL	QDFGPGYSDE	2160
	EPDPGRDEED	LADEMICITT	L				

SEQ ID NO:289 OBI6 DNA SEQUENCE

Nucleic Acid Accession#: NM_002812

Coding sequence: 150-3362 (underlined sequence corresponds to start and stop codon)

20	1	11	21	31	41	51	
	AACTCCCGCC	TCGGGACGCC	TCGGGGTCGG	GCTCCGGCTG	CGGCTGCTGC	TGCGGCGCCC	60
25	CGCGTCCGGT	GCCTCGGCCT	CCTGTGCCCG	CCGCGGAGCA	GTCTGCGGCC	CGCGGTGCGC	120
	CCTCAGCTCC	TTTTCCTGAG	CCGCGCCGGA	TGGGAGCTGC	CGGGGATACC	CCGGCCAGAC	180
	CCGCGCCGGT	GCCTCTGCTC	AGCGTCTGCT	TGCTGCCGCT	GCTGGGCGGT	ACCCAGACAG	240
	CCATTGTCTT	CATCAAGCAG	CCGTCTCTCC	AGGATGCAC	GCAGGGGCGC	CGGGCGCTGC	300
30	TTGCTGTGA	GGTTGAGGCT	CCGGGCCCGG	TACATGTGTA	CTGGCTGCTC	GATGGGGCCC	360
	CTGTCCAGGA	CACGGAGCGG	CGTTTCGCCC	AGGGCAGCAG	CCTGAGCTTT	GCAGCTGTGG	420
	ACCGGCTGCA	GGACTCTGGC	ACCTTCCAGT	GTGTGGCTCG	GGATGATGTC	ACTGGAGAAG	480
	AAGCCCCGAC	TGCCAACGCC	TCCTTCAACA	TCAAATGGAT	TGAGGCAGGT	CCTGTGGTCC	540
	TGAAGCATCC	AGCCTCGGAA	GCTGAGATCC	AGCCACAGAC	CCAGGTACAC	CTTCGTGGCC	600
35	ACATTGATGG	GCACCCCTCG	CCCACCTACC	AATGGTTCCG	AGATGGGACC	CCCCTTTCTG	660
	ATGGTCAGAG	CAACCAACAC	GTCCAGCAGC	AGGAGCGGAA	CCTGACGCTC	CGGCCAGCTG	720
	GTCTGTAGCA	TAGTGGGCTG	TATTCCTGCT	GCGCCACAG	TGCTTTTGGC	CAGGCTTGCA	780
	GCAGCCAGCT	CTTCACCTTG	AGCATGTGCT	ATGAAAGCTT	TGCCAGGGTG	GTGCTGGCAC	840
	CCCAGGACGT	GGTAGTAGCG	AGGTATGAGG	AGGCCATGTT	CCATTGCCAG	TTCTCAGCCC	900
40	AGCCACCCCC	GAGCCTGCAG	TGGCTCTTTG	AGGATGAGAC	TCCCATCACT	AACCCAGATC	960
	CGCCCCCACA	CCTCCGCAGA	GCCACAGTGT	TTGCCAACGG	GTCTCTGCTG	CTGACCCAGG	1020
	TCCGGCCACG	CAATGGGCTG	ATCTACCGCT	GCATTGGCCA	GGGGCAGAGG	GGCCCAACCA	1080
	TCATCTTGGA	AGCCACACTT	CACCTAGCAG	AGATTGAAGA	CATGCCGCTA	TTTGAGCCAC	1140
	GGGTGTTTAC	AGCTGGCAGC	GAGGAGCGTG	TGACCTGCCT	TCCCCCAAG	GGTCTGCCAG	1200
45	AGCCAGCGGT	GTGGTGGGAG	CACGCGGGAG	TCCGGCTGCC	CACCCATGGC	AGGGTCTACC	1260
	AGAAGGGCCA	CGAGCTGGTG	TTGGCCAATA	TTGCTGAAAG	TGATGCTGGT	GTCTACACCT	1320
	GCCACGCGGC	CAACCTGGCT	GGTCAGCGGA	GACAGGATGT	CAACATCACT	GTGGCCACTG	1380
	TGCCCTCCTG	GCTGAAGAAG	CCCCAAGACA	GCCAGCTGGA	GGAGGGCAAA	CCCGGCTACT	1440
	TGGATTGCC	GACCCAGGCG	ACACCAAAAC	CTACAGTTGT	CTGGTACAGA	AACAGATGTC	1500
50	TCATCTCAGA	GGACTCACGG	TTCGAGGTCT	TCAAGAATGG	GACCTTGCGC	ATCAACAGCG	1560
	TGGAGGTGTA	TGATGGGACA	TGGTACCGTT	GTATGAGCAG	CACCCAGACC	GGCAGCATCG	1620
	AGGCGCAAGC	CCGTGTCCAA	GTGCTGGA	AGCTCAAGTT	CACACCACCA	CCCCAGCCAC	1680
	AGCAGTGCAT	GGAGTTTGAC	AAGGAGGCCA	CGGTGCCCTG	TTCAGCCACA	GGCCGAGAGA	1740
	AGCCCACTAT	TAAGTGGGAA	CGGGCAGATG	GGAGCAGCCT	CCCAGAGTGG	GTGACAGACA	1800
55	ACGCTGGGAC	CCTGCATTTT	GCCCGGGTGA	CTCGAGATGA	CGCTGGCAAC	TACACTTGCA	1860
	TTGCTCCAA	CGGGCCGAG	GGCCAGATTC	GTGCCCATGT	CCAGCTCACT	GTGGCAGTTT	1920
	TTATCACCTT	CAAAGTGGAA	CCAGAGCGTA	CGACTGTGTA	CCAGGGCCAC	ACAGCCCTAC	1980
	TGCACTGCCA	GGCCCAAGGG	GACCCCAAGC	CGCTGATICA	GTGGAAAGGC	AAGGACCGCA	2040
60	TCCTGGACCC	CACCAAGCTG	GGACCCAGGA	TGCACATCTT	CCAGAATGGC	TCCCTGGTGA	2100
	TCCATGACGT	GGCCCTGAG	GACTCAGGCC	GCTACACCTG	CATTGCAGGC	AACAGCTGCA	2160
	ACATCAAGCA	CACGGAGGCC	CCCCTCTATG	TCGTGGACAA	GCCTGTGCCG	GAGGAGTCGG	2220
	AGGGCCCTGG	CAGCCCTCCC	CCCTACAAGA	TGATCCAGAC	CATTGGGTTG	TCGGTGGGTTG	2280
	CCGCTGTGGC	CTACATCATAT	GCCGTGCTGG	GCCTCATGTT	CTACTGCAAG	AAGCGCTGCA	2340
	AAGCCAAGCG	GCTGCAGAAG	CAGCCCGAGG	GCGAGGAGCC	AGAGATGGAA	TGCCCTCAACG	2400
65	GAGGGCCCTT	GCAGAACGGG	CAGCCCTCAG	CAGAGATCCA	AGAAGAAGTG	GCCTTGACCA	2460
	GCTTGGGCTC	CGGCCCGCGG	GCCACCAACA	AACGCCACAG	CACAAGTGAT	AAGATGCACT	2520
	TCCCAAGGTC	TAGCCTGCAG	CCCATCACCA	CGCTGGGGAA	GAGTGAGTTT	GGGGAGGTGT	2580
	TCCTGGCAAA	GGCTCAGGGC	TTGGAGGAGG	GAGTGGCAGA	GACCCCTGGTA	CTTGTAAGA	2640
	GCCTGCAGAC	GAAGGATGAG	CAGCAGCAGC	TGGACTTCCG	GAGGGAGTTG	GAGATGTTTG	2700
70	GGAAAGCTGAA	CCACGCCAAC	GTGGTGCGGC	TCCTGGGGCT	GTGCCGGGAG	GCTGAGCCCC	2760
	ACTACATGGT	GCTGGAAATAT	GTGGATCTGG	GAGACCTCAA	GCAGTTCCTG	AGGATTTCCA	2820
	AGAGCAAGGA	TGAAAATATG	AAGTCACAGC	CCCTCAGCAC	CAAGCAGAAG	GTGGCCCTAT	2880
	GCACCCAGGT	AGCCCTGGGG	ATGGAGCACC	TGTCCAACAA	CCGCTTTGTG	CATAAGGACT	2940
	TGGCTGCGCG	TAAGTGCCTG	GTCAGTGCCC	AGAGACAAGT	GAAGGTGTCT	GCCCTGGGCC	3000
	TCAGCAAGGA	TGTGTACAAC	AGTGAGTACT	ACCACTTCCG	CCAGGCCTGG	GTGCCGCTGC	3060
75	GCTGGAATGC	CCCCGAGGCC	ATCCTGGAGG	GTGACTTCTC	TACCAAGTCT	GATGCTTGGG	3120
	CTTTCGGTGT	GCTGATGTGG	GAAGTGTTTA	CACATGGAGA	GATGCCCCAT	GGTGGGCAGG	3180
	CAGATGATGA	AGTACTGGCA	GATTTGCAGG	CTGGGAAGGC	TAGACTTTCCT	CAGCCCGAGG	3240
	GCTGCCCTCT	CAAACTCTAT	CGGCTGATGC	AGCGCTGCTG	GGCCCTCAGC	CCCAAGGACC	3300
80	GGCCCTCCTT	CAGTGAGATT	GCCAGCGCCC	TGGGAGACAG	CACCGTGGAC	AGCAAGCCGT	3360
	GAGGAGGGAG	CCCGCTCAGG	ATGGCCTGGG	CAGGGGAGGA	CATCTCTAGA	GGGAAGCTCA	3420

5	CAGCATGATG	GGCAAGATCC	CTGTCTCTCT	GGGCCCTGAG	GTGCCCTAGT	GCAACAGGCA	3480
	TTGCTGAGGT	CTGAGCAGGG	CCTGGCCCTTT	CCTCCTCTTC	CTCACCCCTCA	TCCTTTGGGA	3540
	GGCTGACTTG	GACCCAAACT	GGGCGACTAG	GGCTTTGAGC	TGGGCAGTTT	CCCCTGCCAC	3600
	CTCTCTCTCT	ATCAGGGACA	GTGTGGGTGC	CACAGGTAAC	CCCAATTTCCT	GGCCTTCAAC	3660
	TTCTCCCTCT	GACCGGGTCC	AACTCTGCCA	CTCATCTGCC	AACTTTGCCT	GGGGAGGGCT	3720
	AGGCTTGGGA	TGAGCTGGGT	TTGTGGGGAG	TTCTTAATA	TTCTCAAGT	CTGGGCACAC	3780
	AGGGTTAATG	AGTCTCTTGC	CCACTGGTCC	ACTTGGGGGT	CTAGACCAGG	ATTATAGAGG	3840
	ACACAGCAAG	TGAGTCTCTCC	CCACTCTGGG	CTTGTGCACA	CTGACCCAGA	CCCACGCTCT	3900
10	CCCCACCTCT	CTCTCTCTTC	CTCATCTTAA	GTGCTGGGCA	GATGAAGGAG	TTTTCAGGAG	3960
	CTTTTTCAC	TATATAAACC	GCCTTTTTCG	TATGCACAC	GGCGCGCTTT	TATATGTAAT	4020
	TGCAGCGTGG	GGTGGGTGGG	CATGGGAGGT	AGGGGTGGGC	CCTGGAGATG	AGGAGGGTGG	4080
	GCCATCTCTA	CCCCACACTT	TTATTGTGTG	CGTTTGTGTG	TTGTTTGTGT	TTTTTGTGTT	4140
	TGTTTTTGTG	TTTACACTCG	CTGCTCTCAA	TAAATAAGCC	TTTTTTTA		

SEQ ID NO:290 QBI6 Protein sequence:

Protein Accession #: NP_002812

20	1	11	21	31	41	51	
	MGAARGSPAR	PRRLPLLSVL	LLPLLGGTQT	AIVFIKQPSS	QDALQRRAL	LRCEVEAPGP	60
	VHYVWLLDGA	PVQDTERRFA	QSSLSFAAV	DRLQDSGTFO	CVARDDVTGE	EARSANASFN	120
	IKWIEAGFVV	LKHPASEAEI	QPQTQVTLRC	HIDGHPRTPT	QWFRDGTPLS	DGQSNHTVSS	180
25	KERNLTLRPA	GPEHSGLYSC	CAHSAFGQAC	SSQNFSLISA	DESFAVVLA	PQDVVVARYE	240
	EAMFHCQFSA	QPPPSLQWLF	EDETPTITNR	RPPHLRRATV	FANGSLLLTQ	VRPRNAGIYR	300
	CIGQQQRGPP	IIIEALTLHLA	EIEDMPLFEP	RVFTAGSEER	VTCLPPKGLP	EPSVWWEHAG	360
	VRLPFTHGRVY	QKQHELVLAN	IAESDAGVYT	CHAANLAGQR	RQDVNITVAT	VPSWLKPKQD	420
	SQLEEGKPGY	LDCLTQATPK	PTVVWYRNQM	LISEDSEFV	FKNGTLRLNS	VEVYDGTWYR	480
	CMSTPAGSI	EAQARQVLE	KLKFTPPPQP	QQCMEDKEA	TVPSCATGRE	KPTIKWERAD	540
30	GSSLPEWVTD	NAGTLHFARV	TRDDAGNYTC	IASNGPQQOI	RAHVQLTVAV	FITFKVEPER	600
	TTYVQGGTAL	LQCEAQDDPK	PLIQWKQKDR	ILDPTKLGR	MHIFQNGSLV	IHDVAPEDSG	660
	RYTCIAGNSC	NIKHTEAPLY	VVDKPVPEES	EGPGSPPPYK	MIQTIGLSVG	AAVAYIIAVL	720
	GLMFYCKKRC	KAKRLQKQPE	GEEPEMECLN	GGPLQNGQPS	AEIQEEVALT	SLGSGPAATN	780
35	KRHSTSDKMH	FPRSSLQPIIT	TLGKSEFGEV	FLAKAQGLEE	GVAETLVLVK	SLQTKDEQQQ	840
	LDFRRELEMF	GKLNHANVVR	LLGLCREAEP	HYMVEYVDL	GDCLKQFLRIS	KSKDEKLKSO	900
	PLSTKQKVAL	CTQVALGMEH	LSNNRFVHKD	LAARNCLVSA	QRQVKVSALG	LSKDVYNSEY	960
	YHFRQAWPVL	RWMSPEAILE	GDFSTKSDVW	AFGVLMWEVF	THGEMPHGGQ	ADDEVLAIDLQ	1020
	AGKARLPQPE	GCPSKLYRLM	QRCWALSPKD	RPSFSEIASA	LGDSTVDSKP		

SEQ ID NO:291 AAB1 DNA SEQUENCE

Nucleic Acid Accession #: NM_002205

Coding sequence:

1-3150 (underlined sequences correspond to start and stop codons)

45	1	11	21	31	41	51	
	ATGGGGAGCC	GGACGCCAGA	GTCCCTCTCT	CACGCCGTGC	AGCTGCGCTG	GGGCCCCCGG	60
50	CGCCGACCCC	CGCTSSGTGC	GCTGCTGTTC	CTGCTSSGTG	CGCCGCCACC	CAGGGTCGGG	120
	GGCTTCAACT	TAGACGCGGA	GGCCCCAGCA	GTACTCTCGG	GGCCCCCGGG	CTCCTTCTTC	180
	GGATTCTCAG	TGGAGTTTTA	CCGGCCGGGA	ACAGACGGGG	TCAGTGTGCT	GGTGGGAGCA	240
	CCCAAGGCTA	ATACCAGCCA	GCCAGGAGTG	CTGCAGGGTG	GTGCTGTCTA	CCTCTGTCTC	300
	TGGGGTGCCA	GCCCCACACA	GTGCACCCCC	ATTGAATTTC	ACAGCAAAGG	CTCTCGGCTC	360
55	CTGGAGTCCT	CACTGTCCAG	CTCAGAGGGA	GAGGAGCCTG	TGGAGTACAA	GTCCTTGCAG	420
	TGGTTCGGGG	CAACAGTTTC	AGCCCATGGC	TCCTCCATCT	TGGCATGCGC	TCCACTGTAC	480
	AGCTGGCGCA	CAGAGAAGGA	GCCACTGAGC	GACCCCGTGG	GCACCTGCTA	CCTCTCCACA	540
	GATAAATTCA	CCCGAATTC	GGAGTATGCA	CCCTGCCGCT	CAGATTTCAG	CTGGGCAGCA	600
	GGACAGGGTT	ACTGCCAAGG	AGGCTTCAGT	GCCGAGTTCA	CCAAGACTGG	CCGTGTGGTT	660
60	TTAGGTGGAC	CAGGAAGCTA	TTTCTGGCAA	GGCCAGATCC	TGTCGTCCAC	TCAGGAGCAG	720
	ATTCAGAAAT	CTTATTACCC	CGAGTACCTG	ATCAACCTGG	TTCAGGGGCA	GCTGCAGACT	780
	CGCCAGGCCA	GTTCCTACTA	TGATGACAGC	TACCTAGGAT	ACTCTGTGGC	TGTTGGTGAA	840
	TTCACTGGTG	ATGACACAGA	AGACTTTGTT	GCTGGTGTGC	CCAAAGGGAA	CCTCACTTAC	900
	GGCTATGTCA	CCATCCTTAA	TGGCTCAGAC	ATTTCGATCCC	TCTACAACCT	CTCAGGGGAA	960
65	CAGATGGCCT	CCTACTTTGG	CTATGCAGTG	GCCGCCACAG	ACGTCAATGG	GGACGGGCTG	1020
	GATGACTTGC	TGGTGGGGGC	ACCCCTGCTC	ATGGATCGGA	CCCCTGACGG	GCGGCCTCAG	1080
	GAGGTGGGCA	GGGTCTACGT	CTACCTGCAG	CACCCAGCCG	GCATAGAGCC	CACGCCACAC	1140
	CTTACCTTCA	CTGGCCATGA	TGAGTTTGGC	CGATTTCGCA	GCTCCTTGAC	CCCCCTGGGG	1200
	GACCTGGACC	AGGATGGCTA	CAATGATGTG	GCCATCGGGG	CTCCCTTTGG	TGGGGAGACC	1260
70	CAGCAGGGAG	TAGTGTTTGT	ATTTCCTGGG	GGCCAGGAG	GGCTGGGCTC	TAAGCCTTCT	1320
	CAGGTTCTGC	AGCCCTGTGT	GGCAGCCAGC	CACACCCAG	ACTTCTTTGG	CTCTGCCCTT	1380
	CGAGGAGGCC	GAGACCTTGA	TGGCAATGGA	TATCCTGATC	TGATTTGTGG	GTCCTTTGGT	1440
	GTGGACAAGG	CTGTGGTATA	CAGGGGCCGC	CCCATCGTGT	CCGCTAGTGC	CTCCCTCACC	1500
	ATCTTCCCCG	CCATGTTTCA	CCCAGAGGAG	CGGAGCTGCA	GCTTAGAGGG	GAACCTGTG	1560
75	GCCTGCATCA	ACCTTAGCTT	CTGCCCTCAAT	GCTTCTGGAA	AACACGTTGC	TGACTCCATT	1620
	GGTTTCACAG	TGGAACCTCA	GCTGGACTGG	CAGAAGCAGA	AGGGAGGGGT	ACGGCGGGCA	1680
	CTGTTCCTGG	CCTCCAGGCA	GGCAACCCCT	ACCCAGACCC	TGCTCATCCA	GAATGGGGCT	1740
	CGAGAGGATT	GCAGAGAGAT	GAAGATCTAC	CTCAGGAACG	AGTCAGAAAT	TCGAGACAAA	1800
	CTCTCGCCGA	TTACATCCGC	TCTCAACTTC	TCCTTGGACC	CCCAAGCCCC	AGTGGACAGC	1860
80	CACGGCCTCA	GCGCCCTCAT	ACATTATCAG	AGCAAGAGCC	GGATAGAGGA	CAAGGCTCAG	1920
	ATCTTGTCTG	ACTGTGGAGA	AGACAACATC	TGTGTGCCCT	ACCTGCAGCT	GGAAGTGTTC	1980

5 GGGGAGCAGA ACCATGTGTA CCTGGGTGAC AAGAATGCCC TGAACCTCAC TTTCCATGCC 2040
 CAGAATGTGG GTGAGGGTGG CGCCTATGAG GCTGAGCTTC GGGTCACCGC CCCTCCAGAG 2100
 GCTGAGTACT CAGGACTCGT CAGACACCCA GGGAACTTCT CCAGCCTGAG CTGTGACTAC 2160
 TTTGCCGTGA ACCAGAGCCG CCTGCTGGTG TGTGACCTGG GCAACCCCAT GAAGGCAGGA 2220
 GCCAGTCTGT GGGGTGGCCT TCGGTTTACA GTCCCTCATC TCCGGGACAC TAAGAAAACC 2280
 ATCCAGTTTG ACTTCCAGAT CCTCAGCAAG AATCTCAACA ACTCGCAAAG CGACGTGGTT 2340
 TCCTTTCGGC TCTCCGTGGA GGCTCAGGCC CAGGTCACCC TGAACGGTGT CTCCAAGCCT 2400
 GAGGCAGTGC TATTTCCAGT AAGCGACTGG CATCCCCGAG ACCAGCCTCA GAAGGAGGAG 2460
 10 GACCTGGGAC CTGCTGTCCA CCATGTCTAT GAGCTCATCA ACCAAGGCCC CAGCTCCATT 2520
 AGCCAGGGTG TGCTGGAAC T CAGCTGTCCC CAGGCTCTGG AAGGTCAGCA GCTCCTATAT 2580
 GTGACCAGAG TTACGGGACT CAACTGCACC ACCAATCACC CCATTAAACC AAAGGGCCTG 2640
 GAGTTGGATC CCGAGGGTTC CCTGCACCAC CAGCAAAAAC GGGAGCTCC AAGCCGCAGC 2700
 TCTGTCTCCT CGGGACCTCA GATCTGAAA TGCCCGGAGG CTGAGTGTTC CAGGCTGCGC 2760
 TGTGAGCTCG GGCCCTTGCA CCAACAAGAG AGCCAAAGTC TGCAAGTTGCA TTTCCGAGTC 2820
 15 TGGGCCAAGA CTTCCTGCA GCGGGAGCAC CAGCCATTTA GCCTGCAGTG TGAGGCTGTG 2880
 TACAAAGCCC TGAAGATGCC CTACCGAATC CTGCTCTGGC AGCTGCCCCA AAAAGAGCGT 2940
 CAGGTGGCCA CAGCTGTGCA ATGGACCAAG GCAGAAGGCA GCTATGGCGT CCCACTGTGG 3000
 ATCATCATCC TAGCCATCCT GTTTGGCCTC CTGCTCCTAG GTCTACTCAT CTACATCCTC 3060
 20 TACAAGCTTG GATTCTTCAA ACGTCCCTC CCATATGGCA CCGCATGGA AAAAGCTCAG 3120
 CTCAAGCCTC CAGCCACCTC TGATGCCTGA

SEQ ID NO:292 AAB1 Protein sequence:

Protein Accession #: NP_002196

25 1 11 21 31 41 51
 MGSRTFESPL HAVQLRWGPR RRPPLPLLL LLLPPPPRVG GFNLDAEAPA VLSGPPGSFF 60
 30 GFSVEFYRFG TDGVSVLVGA PKANTSQPGV LQGGAVYLCF WGASPTQCTP IEFDSKGSRL 120
 LESSLSSEEG BEPVEYKSLQ WFGATVRAHG SSILACAPLY SWRTEKEPLS DPGVTCYLST 180
 DNFTRILEYA PCRSDFSWAA GQGYCQGGFS AEFKTGTGRV LGGPGSYFWQ GQILSATQEQ 240
 IAESYYPEYL INLVQGQLQT RQASSIYDDS YLGYSVAUGE FSGDDTDFV AGVPKGNLTY 300
 GYVITLNGSD IRSLYNFSGE QMASYFGYAV AATDVNGDGL DDLLVGAPLL MDRTPDGRPQ 360
 35 EVGRVYVYLQ HPAGIEPTPT LTLTGHDEFG RFGSSLTPLG DLLDQDGYNDV AIGAPFGGET 420
 QQGVVVFVPG GPGGLGSKPS QVLQPLWAAS HTPDFFGSAL RGGRLDNGG YPDLIVGSFG 480
 VDKAVVYRGR PIVSASASLT IFPAMFNPEE RSCSLEGNFV ACINLSFCLN ASGKHVADSI 540
 GFTVELQLDW QKQKGGVRRR LFLASRQATL TQTLIIQNGA REDCREMKIY LRNESEFRDK 600
 LSPHIALNF SLDPAQAPVDS HGLRPALHYQ SKSRIEDKAO ILLDCGEDNI CVPDLQLEVF 660
 40 GEQNHVYLG D KNALNLTFAH QNVGEGGAYE AELRVATPPE AEYSGLVRHP GNFSLSLSCDY 720
 FAVNQSRLLV CDLGNPMKAG ASLWGLRFT VPHLRDTKKT IQPDFQLSK NLNNSQSDV 780
 SFRLSVEAQA QVTILNGVSKP EAVLFPVSDW HPRDQPKQEE DLGPAVHHVY ELINQGFSSI 840
 SQGVLELSCP QALBGQQLLY VTRVTGLNCT TNHPINPKGL ELDPEGLSHH QQKREAPSR 900
 SASSGPQLLK CPEAEFCFLR CELGPLHQQE SQSLQLHFRV WAKTFLOREH QPFSLQCEAV 960
 45 YKALKMPYRI LPRQLPQKER QVATAVQWTK AEGSYGVFLW IILAILFLGL LLLGLLIYIL 1020
 YKLGFKKRSL PYGTAMEKAQ LKPPATSDA

SEQ ID NO:293 LBH4 DNA SEQUENCE

Nucleic Acid Accession #: BC001291

Coding sequence: 44-541 (start and stop codons are underlined)

50 1 11 21 31 41 51
 55 GGGGGCGCCG CGCGCTGACC CTCCTGGGC ACCGCTGGGG ACGATGGCGC TGCTCGCCTT 60
 GCTGTGGTC GTGGCCCTAC CGCGGGTGTG GACAGACGCC AACCTGACTG CGAGACAACG 120
 AGATCCAGAG GACTCCAGC GAACGGACGA GGGTGACAAT AGAGTGTGGT GTCATGTTG 180
 TGAGAGAGAA AACACTTTCG AGTGCCAGAA CCAAGGAGG TGCAAATGGA CAGAGCCATA 240
 60 CTGCGTTATA GCGGCCGTGA AAATATTTC ACGTTTTTC ATGGTTGCCA AGCAGTGCTC 300
 CGCTGGTGTG GCAGCGATGG AGAGACCCAA GCCAGAGGAG AAGCGGTTTC TCCTGGAAGA 360
 GCCCATGCCC TCTTTTACC TCAAGTGTG TAAAATTCGC TACTGCAATT TAGAGGGGCC 420
 ACCTATCAAC TCATCAGTGT TCAAAGAATA TGCTGGGAGC ATGGGTGAGA GCTGTGGTGG 480
 GCTGTGGCTG GCCATCCTCC TGCTGCTGGC CTCCATTGCA GCCGGCCTCA GCCTGTCTTG 540
 65 AGCCACGGGA CTGCCACAGA CTGAGCCTTC CGGAGCATGG ACTCGCTCCA GACCGTTGTC 600
 ACCTGTTGCA TTAAACTTGT TTTCTGTTGA TTACCTCTTG GTTTGACTTC CCAGGGTCTT 660
 GGGATGGGAG AGTGGGGATC AGGTGCAATT GGCTCTTAAC CCTCAAGGGT TCTTTAACTC 720
 ACATTCAAGG GAAGTCCAGA TCTCCTGAGT AGTGATTTTG GTGACAAGTT TTCTCTTTG 780
 AAATCAAACC TTGTAACCTA TTATTGCTG ATGGCCACTC TTTTCTTGA CTCCCTCTG 840
 70 CCTCTGAGGG CTTCAGTATT GATGGGGAGG GAGGCCTAAG TACCCTCAT GGAGAGTATG 900
 TGCTGAGATG CTTCGACCT TACAGTGAC GCAGGAACAC TGGGGGAGTC TGAATGATTG 960
 GGGTGAAGAC ATCCCTGGAG TGAAGGACTC CTCAGCATGG GGGGCACTGG GGCACACGTT 1020
 AGGGCTGCCC CCAATCCAGT GGTGGAGGCG CTGTGGATGG CTGCTTTTCC TCAACCTTTC 1080
 CTACCAGATT CCAGGAGGCA GAAGATAACT AATTGTGTTG AAGAACTTA GACTTCACCC 1140
 75 ACCAGCTGGC ACAGGTGCAC AGATTCTATA ATTCCACAC GTGTGTGTTT AACATCTGAA 1200
 ACTTAGGCCA AGTAGAGAGC ATCAGGGTAA ATGGCGTTCA TTTCTCTGTT AAGATGCAGC 1260
 CATCCATGGG GAGCTGAGAA ATCAGACTCA AAGTCCACC AAAAACAAT ACAAGGGGAC 1320
 TTCAAAAGTT CACGAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAA

SEQ ID NO:294 LBH4 Protein sequence:
Protein Accession #: AAH01291

5 1 11 21 31 41 51
| | | | | |
MALLALLLVV ALPRVWTDAN LTARQRPED SQRTEGDNR VWCHVCEREN TFECQNP RRC 60
10 KWTEPYCVIA AVKIFPRFFM VAKQCSAGCA AMERPKPEEK RFLLEPMPF FYLKCCCKIRY 120
CNLEGPPINS SVFKEYAGSM GESCGGLWLA ILLLLASIAA GLSLS

15 It is understood that the examples described above in no way serve to limit the
true scope of this invention, but rather are presented for illustrative purposes. All
publications, sequences of accession numbers, and patent applications cited in this
specification are herein incorporated by reference as if each individual publication or patent
20 application were specifically and individually indicated to be incorporated by reference.

WHAT IS CLAIMED IS:

- 1 1. A method of detecting a prostate cancer-associated transcript in a cell
2 from a patient, the method comprising contacting a biological sample from the patient with a
3 polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence
4 as shown in Tables 1-16.
- 1 2. The method of claim 1, wherein the polynucleotide selectively
2 hybridizes to a sequence at least 95% identical to a sequence as shown in Tables 1-16.
- 1 3. The method of claim 1, wherein the biological sample is a tissue
2 sample.
- 1 4. The method of claim 1, wherein the biological sample comprises
2 isolated nucleic acids.
- 1 5. The method of claim 4, wherein the nucleic acids are mRNA.
- 1 6. The method of claim 4, further comprising the step of amplifying
2 nucleic acids before the step of contacting the biological sample with the polynucleotide.
- 1 7. The method of claim 1, wherein the polynucleotide comprises a
2 sequence as shown in Tables 1-16.
- 1 8. The method of claim 1, wherein the polynucleotide is labeled.
- 1 9. The method of claim 8, wherein the label is a fluorescent label.
- 1 10. The method of claim 1, wherein the polynucleotide is immobilized on
2 a solid surface.
- 1 11. The method of claim 1, wherein the patient is undergoing a therapeutic
2 regimen to treat prostate cancer.
- 1 12. The method of claim 1, wherein the patient is suspected of having
2 prostate cancer.

1 13. A method of monitoring the efficacy of a therapeutic treatment of
2 prostate cancer, the method comprising the steps of:

3 (i) providing a biological sample from a patient undergoing the therapeutic
4 treatment; and

5 (ii) determining the level of a prostate cancer-associated transcript in the
6 biological sample by contacting the biological sample with a polynucleotide that selectively
7 hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-16,
8 thereby monitoring the efficacy of the therapy.

1 14. The method of claim 13, further comprising the step of: (iii) comparing
2 the level of the prostate cancer-associated transcript to a level of the prostate cancer-
3 associated transcript in a biological sample from the patient prior to, or earlier in, the
4 therapeutic treatment.

1 15. The method of claim 13, wherein the patient is a human.

1 16. A method of monitoring the efficacy of a therapeutic treatment of
2 prostate cancer, the method comprising the steps of:

3 (i) providing a biological sample from a patient undergoing the therapeutic
4 treatment; and

5 (ii) determining the level of a prostate cancer-associated antibody in the
6 biological sample by contacting the biological sample with a polypeptide encoded by a
7 polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence
8 as shown in Tables 1-16, wherein the polypeptide specifically binds to the prostate cancer-
9 associated antibody, thereby monitoring the efficacy of the therapy.

1 17. The method of claim 16, further comprising the step of: (iii) comparing
2 the level of the prostate cancer-associated antibody to a level of the prostate cancer-
3 associated antibody in a biological sample from the patient prior to, or earlier in, the
4 therapeutic treatment.

1 18. The method of claim 16, wherein the patient is a human.

1 19. A method of monitoring the efficacy of a therapeutic treatment of
2 prostate cancer, the method comprising the steps of:

3 (i) providing a biological sample from a patient undergoing the therapeutic
4 treatment; and

5 (ii) determining the level of a prostate cancer-associated polypeptide in the
6 biological sample by contacting the biological sample with an antibody, wherein the antibody
7 specifically binds to a polypeptide encoded by a polynucleotide that selectively hybridizes to
8 a sequence at least 80% identical to a sequence as shown in Tables 1-16, thereby monitoring
9 the efficacy of the therapy.

1 20. The method of claim 19, further comprising the step of: (iii) comparing
2 the level of the prostate cancer-associated polypeptide to a level of the prostate cancer-
3 associated polypeptide in a biological sample from the patient prior to, or earlier in, the
4 therapeutic treatment.

1 21. The method of claim 19, wherein the patient is a human.

1 22. An isolated nucleic acid molecule consisting of a polynucleotide
2 sequence as shown in Tables 1-16.

1 23. The nucleic acid molecule of claim 22, which is labeled.

1 24. The nucleic acid of claim 23, wherein the label is a fluorescent label

1 25. An expression vector comprising the nucleic acid of claim 22.

1 26. A host cell comprising the expression vector of claim 25.

1 27. An isolated polypeptide which is encoded by a nucleic acid molecule
2 having polynucleotide sequence as shown in Tables 1-16.

1 28. An antibody that specifically binds a polypeptide of claim 27.

1 29. The antibody of claim 28, further conjugated to an effector component.

1 30. The antibody of claim 29, wherein the effector component is a
2 fluorescent label.

1 31. The antibody of claim 29, wherein the effector component is a
2 radioisotope or a cytotoxic chemical.

1 32. The antibody of claim 29, which is an antibody fragment.

1 33. The antibody of claim 29, which is a humanized antibody

1 34. A method of detecting a prostate cancer cell in a biological sample
2 from a patient, the method comprising contacting the biological sample with an antibody of
3 claim 28.

1 35. The method of claim 34, wherein the antibody is further conjugated to
2 an effector component.

1 36. The method of claim 35, wherein the effector component is a
2 fluorescent label.

1 37. A method of detecting antibodies specific to prostate cancer in a
2 patient, the method comprising contacting a biological sample from the patient with a
3 polypeptide encoded by a nucleic acid comprises a sequence from Tables 1-16.

1 38. A method for identifying a compound that modulates a prostate cancer-
2 associated polypeptide, the method comprising the steps of:

3 (i) contacting the compound with a prostate cancer-associated polypeptide, the
4 polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least
5 80% identical to a sequence as shown in Tables 1-16; and

6 (ii) determining the functional effect of the compound upon the polypeptide.

1 39. The method of claim 38, wherein the functional effect is a physical
2 effect.

1 40. The method of claim 38, wherein the functional effect is a chemical
2 effect.

1 41. The method of claim 38, wherein the polypeptide is expressed in a
2 eukaryotic host cell or cell membrane.

1 42. The method of claim 38, wherein the functional effect is determined by
2 measuring ligand binding to the polypeptide.

1 43. The method of claim 38, wherein the polypeptide is recombinant.

1 44. A method of inhibiting proliferation of a prostate cancer-associated
2 cell to treat prostate cancer in a patient, the method comprising the step of administering to
3 the subject a therapeutically effective amount of a compound identified using the method of
4 claim 38.

1 45. The method of claim 44, wherein the compound is an antibody.

1 46. The method of claim 45, wherein the patient is a human.

1 47. A drug screening assay comprising the steps of
2 (i) administering a test compound to a mammal having prostate cancer or a
3 cell isolated therefrom;
4 (ii) comparing the level of gene expression of a polynucleotide that selectively
5 hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-16 in a
6 treated cell or mammal with the level of gene expression of the polynucleotide in a control
7 cell or mammal, wherein a test compound that modulates the level of expression of the
8 polynucleotide is a candidate for the treatment of prostate cancer.

1 48. The assay of claim 47, wherein the control is a mammal with prostate
2 cancer or a cell therefrom that has not been treated with the test compound.

1 49. The assay of claim 47, wherein the control is a normal cell or mammal.

1 50. A method for treating a mammal having prostate cancer comprising
2 administering a compound identified by the assay of claim 47.

1 51. A pharmaceutical composition for treating a mammal having prostate
2 cancer, the composition comprising a compound identified by the assay of claim 47 and a
3 physiologically acceptable excipient.

1 52. The method according to claim 1, wherein said biological sample is
2 contacted with a plurality of polynucleotides comprising a first polynucleotide that
3 selectively hybridizes to a sequence at least 80% identical to a first sequence as shown in
4 Tables 1-16; and a second polynucleotide that selectively hybridizes to a second sequence at
5 least 80% identical to a second sequence as shown in Tables 1-16.

1 53. A method according to claim 52, wherein the plurality of
2 polynucleotides comprises a third polynucleotide that selectively hybridizes to a sequence at
3 least 80% identical to a third sequence as shown in Tables 1-16..

1 54. A method of detecting a prostate cancer associated transcript, the
2 method comprising contacting a biological sample from the patient with a plurality of
3 polynucleotides wherein at least two of said polynucleotides selectively hybridize to a
4 difference sequence at least 80% identical to a sequence as shown in Tables 1-16.

1 55. A method of detecting a prostate cancer, the method comprising the
2 steps of:

3 (i) providing a biological sample from a patient;

4 (ii) contacting the biological sample with a first polynucleotide that selectively
5 hybridizes to a sequence at least 80% identical to a first sequence as shown in Tables 1-16 to
6 determine the level of a prostate cancer-associated transcript in the biological sample; and
7 with a second polynucleotide that selectively hybridizes to a second sequence at least 80%
8 identical to a sequence not shown in Tables 1-16; wherein the expression of said second
9 sequence is not substantially changed in prostate cancer, to determine the level of expression
10 of a control transcript in the biological sample;

11 (iii) comparing the level of the prostate cancer-associated transcript to a level
12 of the normal tissue associated transcript in the biological sample.

1 56. A method of quantitating a prostate cancer-associated transcript in a
2 cell from a patient, the method comprising contacting a biological sample from the patient
3 with a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a
4 sequence as shown in Tables 1-16.

1 57. The method of claim 56, wherein the polynucleotide selectively
2 hybridizes to a sequence at least 95% identical to a sequence as shown in Tables 1-16.

1 58. The method of claim 56, wherein the biological sample is a tissue
2 sample.

1 59. The method of claim 56, wherein the biological sample comprises
2 isolated nucleic acids.

1 60. The method of claim 56, wherein the nucleic acids are mRNA.

1 61. The method of claim 59, further comprising the step of amplifying
2 nucleic acids before the step of contacting the biological sample with the polynucleotide.

1 62. The method of claim 56, wherein the polynucleotide comprises a
2 sequence as shown in Tables 1-16.

1 63. The method of claim 56, wherein the polynucleotide is labeled.

1 64. The method of claim 63, wherein the label is a fluorescent label.

1 65. The method of claim 56, wherein the polynucleotide is immobilized on
2 a solid surface.

1 66. The method of claim 56, wherein the patient is undergoing a
2 therapeutic regimen to treat metastatic prostate cancer.

1 67. The method of claim 56, wherein the patient is suspected of having
2 metastatic prostate cancer.

1 68. A biochip comprising a plurality of polynucleotides that selectively
2 hybridize to a sequence at least 80% identical to a sequence as shown in Tables 1-16.

1 69. A method of screening drug candidates comprising:
2 i) providing a cell that expresses an expression profile gene selected from the
3 group consisting of an expression profile gene set forth in Tables 1-16 or fragment thereof;
4 ii) adding a drug candidate to said cell; and
5 iii) determining the effect of said drug candidate on the expression of said
6 expression profile gene.

1 70. A method according to claim 59 wherein said determining comprises
2 comparing the level of expression in the absence of said drug candidate to the level of
3 expression in the presence of said drug candidate.

1 SF 1277890 v1